

特許協力条約に基づく国際出願願書

副本 - 印刷日時 2001年02月16日 (16.02.2001) 金曜日 15時07分50秒

0	受理官庁記入欄	
0-1	国際出願番号.	
0-2	国際出願日	
0-3	(受付印)	
0-4	様式-PCT/R0/101 この特許協力条約に基づく国際出願願書は、 右記によって作成された。	PCT-EASY Version 2.91 (updated 01.01.2001)
0-5	申立て 出願人は、この国際出願が特許協力条約に従って処理されることを請求する。	
0-6	出願人によって指定された受理官庁	日本国特許庁 (R0/JP)
0-7	出願人又は代理人の書類記号	SK01PCT13
I	発明の名称	音声認識装置及び音声認識方法、並びに記録媒体
II	出願人	出願人である (applicant only)
II-1	この欄に記載した者は	米国を除くすべての指定国 (all designated States except US)
II-2	右の指定国についての出願人である。	
II-4ja	名称	ソニー株式会社
II-4en	Name	SONY CORPORATION
II-5ja	あて名:	141-0001 日本国
II-5en	Address:	東京都 品川区 北品川6丁目7番35号 7-35, Kitashinagawa 6-chome Shinagawa-ku, Tokyo 141-0001 Japan
II-6	国籍 (国名)	日本国 JP
II-7	住所 (国名)	日本国 JP
III-1	その他の出願人又は発明者	出願人及び発明者である (applicant and inventor)
III-1-1	この欄に記載した者は	米国のみ (US only)
III-1-2	右の指定国についての出願人である。	
III-1-4ja	氏名 (姓名)	浅野 康治
III-1-4en	Name (LAST, First)	ASANO, Yasuharu
III-1-5ja	あて名:	141-0001 日本国
III-1-5en	Address:	東京都 品川区 北品川6丁目7番35号 ソニー株式会社内 c/o SONY CORPORATION 7-35, Kitashinagawa 6-chome Shinagawa-ku, Tokyo 141-0001 Japan
III-1-6	国籍 (国名)	日本国 JP
III-1-7	住所 (国名)	日本国 JP

THIS PAGE BLANK (USPTO)

特許協力条約に基づく国際出願願書

副本 - 印刷日時 2001年02月16日 (16.02.2001) 金曜日 15時07分50秒

III-2	その他の出願人又は発明者	出願人及び発明者である (applicant and inventor)
III-2-1	この欄に記載した者は	米国のみ (US only)
III-2-2	右の指定国についての出願人である。	
III-2-4ja	氏名(姓名)	南野 活樹
III-2-4en	Name (LAST, First)	MINAMINO, Katsuki
III-2-5ja	あて名:	141-0001 日本国
		東京都 品川区
		北品川 6 丁目 7 番 3 5 号
		ソニー株式会社内
III-2-5en	Address:	c/o SONY CORPORATION
		7-35, Kitashinagawa 6-chome
		Shinagawa-ku, Tokyo 141-0001
		Japan
III-2-6	国籍 (国名)	日本国 JP
III-2-7	住所 (国名)	日本国 JP
III-3	その他の出願人又は発明者	出願人及び発明者である (applicant and inventor)
III-3-1	この欄に記載した者は	米国のみ (US only)
III-3-2	右の指定国についての出願人である。	
III-3-4ja	氏名(姓名)	小川 浩明
III-3-4en	Name (LAST, First)	OGAWA, Hiroaki
III-3-5ja	あて名:	141-0001 日本国
		東京都 品川区
		北品川 6 丁目 7 番 3 5 号
		ソニー株式会社内
III-3-5en	Address:	c/o SONY CORPORATION
		7-35, Kitashinagawa 6-chome
		Shinagawa-ku, Tokyo 141-0001
		Japan
III-3-6	国籍 (国名)	日本国 JP
III-3-7	住所 (国名)	日本国 JP
III-4	その他の出願人又は発明者	出願人及び発明者である (applicant and inventor)
III-4-1	この欄に記載した者は	米国のみ (US only)
III-4-2	右の指定国についての出願人である。	
III-4-4ja	氏名(姓名)	ルッケ ヘルムート
III-4-4en	Name (LAST, First)	LUCKE, Helmut
III-4-5ja	あて名:	141-0001 日本国
		東京都 品川区
		北品川 6 丁目 7 番 3 5 号
		ソニー株式会社内
III-4-5en	Address:	c/o SONY CORPORATION
		7-35, Kitashinagawa 6-chome
		Shinagawa-ku, Tokyo 141-0001
		Japan
III-4-6	国籍 (国名)	ドイツ連邦共和国 DE
III-4-7	住所 (国名)	日本国 JP

THIS PAGE BLANK (USPTO)

特許協力条約に基づく国際出願願書

副本 - 印刷日時 2001年02月16日 (16.02.2001) 金曜日 15時07分50秒

IV-1	代理人又は共通の代表者、通知のあて名 下記の者は国際機関において右記のごとく出願人のために行動する。	代理人 (agent)
IV-1-1ja	氏名(姓名)	小池 晃
IV-1-1en	Name (LAST, First)	KOIKE, Akira
IV-1-2ja	あて名:	105-0001 日本国 東京都 港区 虎ノ門二丁目6番4号 第11森ビル
IV-1-2en	Address:	No.11 Mori Bldg., 6-4, Toranomon 2-chome Minato-ku, Tokyo 105-0001 Japan
IV-1-3	電話番号	03-3508-8266
IV-1-4	ファクシミリ番号	03-3508-0439
IV-2	その他の代理人	筆頭代理人と同じあて名を有する代理人 (additional agent(s) with same address as first named agent)
IV-2-1ja	氏名	田村 榮一; 伊賀 誠司
IV-2-1en	Name(s)	TAMURA, Eiichi; IGA, Seiji
V	国の指定	
V-1	広域特許 (他の種類の保護又は取扱いを求める場合には括弧内に記載する。)	EP: AT BE CH&LI CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR 及びヨーロッパ特許条約と特許協力条約の締約国である他の国
V-2	国内特許 (他の種類の保護又は取扱いを求める場合には括弧内に記載する。)	CN JP US
V-5	指定の確認の宣言 出願人は、上記の指定に加えて、規則4.9(b)の規定に基づき、特許協力条約のもとで認められる他の全ての国の指定を行う。ただし、V-6欄に示した国の指定を除く。出願人は、これらの追加される指定が確認を条件としていること、並びに優先日から15月が経過する前にその確認がなされない指定は、この期間の経過時に、出願人によって取り下げられたものとみなされることを宣言する。	
V-6	指定の確認から除かれる国	なし (NONE)
VI-1	先の国内出願に基づく優先権主張	
VI-1-1	先の出願日	2000年02月28日 (28.02.2000)
VI-1-2	先の出願番号	特願2000-051466
VI-1-3	国名	日本国 JP
VII-1	特定された国際調査機関(ISA)	日本国特許庁 (ISA/JP)

THIS PAGE BLANK (USPTO)

特許協力条約に基づく国際出願願書

副本 - 印刷日時 2001年02月16日 (16.02.2001) 金曜日 15時07分50秒

VIII	照合欄	用紙の枚数	添付された電子データ
VIII-1	願書	5	-
VIII-2	明細書	26	-
VIII-3	請求の範囲	2	-
VIII-4	要約	1	absk01pct13.txt
VIII-5	図面	7	-
VIII-7	合計	41	
	添付書類	添付	添付された電子データ
VIII-8	手数料計算用紙	✓	-
VIII-10	包括委任状の写し	✓	-
VIII-12	優先権証明書	優先権証明書 VI-1	-
VIII-16	PCT-EASYディスク	-	フレキシブルディスク
VIII-17	その他	納付する手数料に相当する特許印紙を貼付した書面	-
VIII-17	その他	国際事務局の口座への振込を証明する書面	-
VIII-18	要約書とともに提示する図の番号	2	
VIII-19	国際出願の使用言語名:	日本語 (Japanese)	
IX-1	提出者の記名押印		
IX-1-1	氏名(姓名)	小池 晃	
IX-2	提出者の記名押印		
IX-2-1	氏名(姓名)	田村 榮一	
IX-3	提出者の記名押印		
IX-3-1	氏名(姓名)	伊賀 誠司	

受理官庁記入欄

10-1	国際出願として提出された書類の実際の受理の日	
10-2	図面:	
10-2-1	受理された	
10-2-2	不足図面がある	
10-3	国際出願として提出された書類を補完する書類又は図面であってその後期間内に提出されたものの実際の受理の日(訂正日)	
10-4	特許協力条約第11条(2)に基づく必要な補完の期間内の受理の日	
10-5	出願人により特定された国際調査機関	ISA/JP

THIS PAGE BLANK (USPTO)

特許協力条約に基づく国際出願願書

SK01PCT13

副本 - 印刷日時 2001年02月16日 (16.02.2001) 金曜日 15時07分50秒

10-6	調査手数料未払いにつき、国際調査機関に調査用写しを送付していない	
------	----------------------------------	--

国際事務局記入欄

11-1	記録原本の受理の日	
------	-----------	--

THIS PAGE BLANK (USPTO)

PATENT COOPERATION TREATY

PCT

From the INTERNATIONAL BUREAU

NOTIFICATION OF RECEIPT OF
RECORD COPY

(PCT Rule 24.2(a))

To:

KOIKE, Akira
No.11 Mori Bldg., 6-4, Toranomon 2-
chome
Minato-ku, Tokyo 105-0001
JAPON

Date of mailing (day/month/year) 20 March 2001 (20.03.01)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference SK01PCT13	International application No. PCT/JP01/01127

The applicant is hereby notified that the International Bureau has received the record copy of the international application as detailed below.

Name(s) of the applicant(s) and State(s) for which they are applicants:

SONY CORPORATION (for all designated States except US)
ASANO, Yasuharu et al (for US)

International filing date : 16 February 2001 (16.02.01)
Priority date(s) claimed : 28 February 2000 (28.02.00)
Date of receipt of the record copy
by the International Bureau : 02 March 2001 (02.03.01)
List of designated Offices :

EP : AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR
National : CN, JP, US

ATTENTION

The applicant should carefully check the data appearing in this Notification. In case of any discrepancy between these data and the indications in the international application, the applicant should immediately inform the International Bureau.

In addition, the applicant's attention is drawn to the information contained in the Annex, relating to:

- ☒ time limits for entry into the national phase
☒ confirmation of precautionary designations
☐ requirements regarding priority documents

A copy of this Notification is being sent to the receiving Office and to the International Searching Authority.

<p>The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland</p> <p>Facsimile No. (41-22) 740.14.35</p>	<p>Authorized officer:</p> <p>H. Zhou</p> <p>Telephone No. (41-22) 338.83.38</p>
---	--

THIS PAGE BLANK (USPTO)

INFORMATION ON TIME LIMITS FOR ENTERING THE NATIONAL PHASE

The applicant is reminded that the "national phase" must be entered before each of the designated Offices indicated in the Notification of Receipt of Record Copy (Form PCT/IB/301) by paying national fees and furnishing translations, as prescribed by the applicable national laws.

The time limit for performing these procedural acts is **20 MONTHS** from the priority date or, for those designated States which the applicant elects in a demand for international preliminary examination or in a later election, **30 MONTHS** from the priority date, provided that the election is made before the expiration of 19 months from the priority date. Some designated (or elected) Offices have fixed time limits which expire even later than 20 or 30 months from the priority date. In other Offices an extension of time or grace period, in some cases upon payment of an additional fee, is available.

In addition to these procedural acts, the applicant may also have to comply with other special requirements applicable in certain Offices. **It is the applicant's responsibility** to ensure that the necessary steps to enter the national phase are taken in a timely fashion. Most designated Offices do not issue reminders to applicants in connection with the entry into the national phase.

For detailed information about the procedural acts to be performed to enter the national phase before each designated Office, the applicable time limits and possible extensions of time or grace periods, and any other requirements, see the relevant Chapters of Volume II of the PCT Applicant's Guide. Information about the requirements for filing a demand for international preliminary examination is set out in Chapter IX of Volume I of the PCT Applicant's Guide.

GR and ES became bound by PCT Chapter II on 7 September 1996 and 6 September 1997, respectively, and may, therefore, be elected in a demand or a later election filed on or after 7 September 1996 and 6 September 1997, respectively, regardless of the filing date of the international application. (See second paragraph above.)

Note that only an applicant who is a national or resident of a PCT Contracting State which is bound by Chapter II has the right to file a demand for international preliminary examination.

CONFIRMATION OF PRECAUTIONARY DESIGNATIONS

This notification lists only specific designations made under Rule 4.9(a) in the request. It is important to check that these designations are correct. Errors in designations can be corrected where precautionary designations have been made under Rule 4.9(b). The applicant is hereby reminded that any precautionary designations may be confirmed according to Rule 4.9(c) before the expiration of 15 months from the priority date. If it is not confirmed, it will automatically be regarded as withdrawn by the applicant. There will be no reminder and no invitation. Confirmation of a designation consists of the filing of a notice specifying the designated State concerned (with an indication of the kind of protection or treatment desired) and the payment of the designation and confirmation fees. Confirmation must reach the receiving Office within the 15-month time limit.

REQUIREMENTS REGARDING PRIORITY DOCUMENTS

For applicants who have not yet complied with the requirements regarding priority documents, the following is recalled.

Where the priority of an earlier national, regional or international application is claimed, the applicant must submit a copy of the said earlier application, certified by the authority with which it was filed ("the priority document") to the receiving Office (which will transmit it to the International Bureau) or directly to the International Bureau, before the expiration of 16 months from the priority date, provided that any such priority document may still be submitted to the International Bureau before that date of international publication of the international application, in which case that document will be considered to have been received by the International Bureau on the last day of the 16-month time limit (Rule 17.1(a)).

Where the priority document is issued by the receiving Office, the applicant may, instead of submitting the priority document, request the receiving Office to prepare and transmit the priority document to the International Bureau. Such request must be made before the expiration of the 16-month time limit and may be subjected by the receiving Office to the payment of a fee (Rule 17.1(b)).

If the priority document concerned is not submitted to the International Bureau or if the request to the receiving Office to prepare and transmit the priority document has not been made (and the corresponding fee, if any, paid) within the applicable time limit indicated under the preceding paragraphs, any designated State may disregard the priority claim, provided that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity to furnish the priority document within a time limit which is reasonable under the circumstances.

Where several priorities are claimed, the priority date to be considered for the purposes of computing the 16-month time limit is the filing date of the earliest application whose priority is claimed.

THIS PAGE BLANK (USPTO)

PATENT COOPERATION TREATY

PCT

NOTIFICATION CONCERNING
SUBMISSION OR TRANSMITTAL
OF PRIORITY DOCUMENT

(PCT Administrative Instructions, Section 411)

From the INTERNATIONAL BUREAU

To:

KOIKE, Akira
No.11 Mori Bldg., 6-4, Toranomom 2-
chome
Minato-ku, Tokyo 105-0001
JAPON

Date of mailing (day/month/year) 20 March 2001 (20.03.01)	
Applicant's or agent's file reference SK01PCT13	IMPORTANT NOTIFICATION
International application No. PCT/JP01/01127	International filing date (day/month/year) 16 February 2001 (16.02.01)
International publication date (day/month/year) Not yet published	Priority date (day/month/year) 28 February 2000 (28.02.00)
Applicant SONY CORPORATION et al	

1. The applicant is hereby notified of the date of receipt (except where the letters "NR" appear in the right-hand column) by the International Bureau of the priority document(s) relating to the earlier application(s) indicated below. Unless otherwise indicated by an asterisk appearing next to a date of receipt, or by the letters "NR", in the right-hand column, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).
2. This updates and replaces any previously issued notification concerning submission or transmittal of priority documents.
3. An asterisk(*) appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b). In such a case, **the attention of the applicant is directed** to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.
4. The letters "NR" appearing in the right-hand column denote a priority document which was not received by the International Bureau or which the applicant did not request the receiving Office to prepare and transmit to the International Bureau, as provided by Rule 17.1(a) or (b), respectively. In such a case, **the attention of the applicant is directed** to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

<u>Priority date</u>	<u>Priority application No.</u>	<u>Country or regional Office or PCT receiving Office</u>	<u>Date of receipt of priority document</u>
28 Febr 2000 (28.02.00)	2000/51466	JP	02 Marc 2001 (02.03.01)

<p>The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland</p> <p>Facsimile No. (41-22) 740.14.35</p>	<p>Authorized officer H. Zhou</p> <p>Telephone No. (41-22) 338.83.38</p>
---	--

THIS PAGE BLANK (USPTO)

PATENT COOPERATION TREATY

PCT

NOTICE INFORMING THE APPLICANT OF THE COMMUNICATION OF THE INTERNATIONAL APPLICATION TO THE DESIGNATED OFFICES

(PCT Rule 47.1(c), first sentence)

From the INTERNATIONAL BUREAU

To:

KOIKE, Akira
No.11 Mori Bldg., 6-4, Toranomon 2-
chome
Minato-ku, Tokyo 105-0001
JAPON

Date of mailing (day/month/year) 06 September 2001 (06.09.01)		
Applicant's or agent's file reference SK01PCT13		IMPORTANT NOTICE
International application No. PCT/JP01/01127	International filing date (day/month/year) 16 February 2001 (16.02.01)	
Priority date (day/month/year) 28 February 2000 (28.02.00)		
Applicant SONY CORPORATION et al		

1. Notice is hereby given that the International Bureau has communicated, as provided in Article 20, the international application to the following designated Offices on the date indicated above as the date of mailing of this Notice:
- US

In accordance with Rule 47.1(c), third sentence, those Offices will accept the present Notice as conclusive evidence that the communication of the international application has duly taken place on the date of mailing indicated above and no copy of the international application is required to be furnished by the applicant to the designated Office(s).

2. The following designated Offices have waived the requirement for such a communication at this time:
- CN,EP,JP

The communication will be made to those Offices only upon their request. Furthermore, those Offices do not require the applicant to furnish a copy of the international application (Rule 49.1(a-bis)).

3. Enclosed with this Notice is a copy of the international application as published by the International Bureau on 07 September 2001 (07.09.01) under No. WO 01/65541

REMINDER REGARDING CHAPTER II (Article 31(2)(a) and Rule 54.2)

If the applicant wishes to postpone entry into the national phase until 30 months (or later in some Offices) from the priority date, a **demand for international preliminary examination** must be filed with the competent International Preliminary Examining Authority before the expiration of 19 months from the priority date.

It is the applicant's sole responsibility to monitor the 19-month time limit.

Note that only an applicant who is a national or resident of a PCT Contracting State which is bound by Chapter II has the right to file a demand for international preliminary examination.

REMINDER REGARDING ENTRY INTO THE NATIONAL PHASE (Article 22 or 39(1))

If the applicant wishes to proceed with the international application in the **national phase**, he must, within 20 months or 30 months, or later in some Offices, perform the acts referred to therein before each designated or elected Office.

For further important information on the time limits and acts to be performed for entering the national phase, see the Annex to Form PCT/IB/301 (Notification of Receipt of Record Copy) and Volume II of the PCT Applicant's Guide.

<p>The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland</p> <p>Facsimile No. (41-22) 740.14.35</p>	<p>Authorized officer</p> <p>J. Zahra</p> <p>Telephone No. (41-22) 338.83.38</p>
---	--

THIS PAGE BLANK (USPTO)

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2001年9月7日 (07.09.2001)

PCT

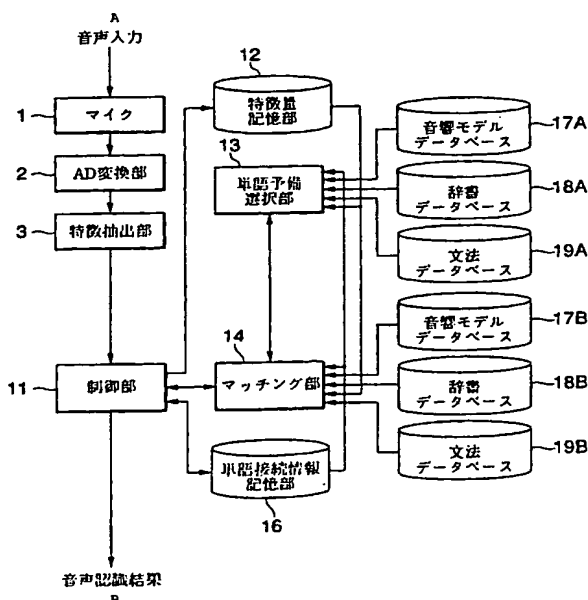
(10) 国際公開番号
WO 01/65541 A1

- (51) 国際特許分類: G10L 15/28, 15/18 (ASANO, Yasuharu) [JP/JP]. 南野活樹 (MINAMINO, Katsuki) [JP/JP]. 小川浩明 (OGAWA, Hiroaki) [JP/JP]. ルッケ ヘルムート (LUCKE, Helmut) [DE/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo (JP).
- (21) 国際出願番号: PCT/JP01/01127
- (22) 国際出願日: 2001年2月16日 (16.02.2001)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ: 特願2000-51466 2000年2月28日 (28.02.2000) JP
- (71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 Tokyo (JP).
- (72) 発明者; および (75) 発明者/出願人 (米国についてのみ): 浅野康治
- (74) 代理人: 小池 晃, 外 (KOIKE, Akira et al.); 〒105-0001 東京都港区虎ノ門二丁目6番4号 第11森ビル Tokyo (JP).
- (81) 指定国 (国内): CN, JP, US.
- (84) 指定国 (広域): ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).
- 添付公開書類:
— 国際調査報告書

[続葉有]

(54) Title: SPEECH RECOGNITION DEVICE AND SPEECH RECOGNITION METHOD, AND RECORDING MEDIUM

(54) 発明の名称: 音声認識装置及び音声認識方法、並びに記録媒体



(57) Abstract: A speech recognition device for improving a speech recognition accuracy while restricting an increase in resource, wherein a probable word as the result of speech recognition is selected based on acoustic scores and linguistic scores, and, in addition, a word is selected based on criteria other than acoustic scores such as "small number of phonemes", "being specified part of speech", "included in the past speech recognition results", and "linguistic score higher than a preset value", and the selected word is subjected to matching.

- A...VOICE INPUT
1...MICROPHONE
2...A/D CONVERTER
3...FEATURE EXTRACTING UNIT
11...CONTROL UNIT
B...SPEECH RECOGNITION RESULT
12...FEATURE QUANTITY STORING UNIT
13...WORD PRE-SELECTING UNIT
14...MATCHING UNIT
16...WORD CONNECTION INFORMATION STORING UNIT
17A...ACOUSTIC MODEL DATABASE
18A...DICTIONARY DATABASE
19A...GRAMMAR DATABASE
17B...ACOUSTIC MODEL DATABASE
18B...DICTIONARY DATABASE
19B...GRAMMAR DATABASE

[続葉有]

WO 01/65541 A1

THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/01127

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl⁷ G10L15/28, 15/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl⁷ G10L15/00-17/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2001
Kokai Jitsuyo Shinan Koho 1971-2001 Jitsuyo Shinan Toroku Koho 1996-2001

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JICST FILE (JOIS)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	IEEE Transactions on Speech and Audio Processing, Vol.1, No.1, January 1993, L. R. Bahl et al., "A Fast Approximate Acoustic Match for Large Vocabulary Speech Recognition", pp.59-67, ISSN1063-6676	1-9
A	JP, 57-86899, A (Mitsubishi Electric Corporation), 31 May, 1982 (31.05.82), Full text, all drawings (Family: none)	1-9
A	JP, 58-52696, A (Dainippon Printing Co., Ltd.), 28 March, 1983 (28.03.83), Full text, all drawings (Family: none)	1-9

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search
29 March, 2001 (29.03.01)

Date of mailing of the international search report
10 April, 2001 (10.04.01)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

THIS PAGE BLANK (USPTO)

E P . U S P C T

国際調査報告

(法 8 条、法施行規則第40、41条)
〔P C T 1 8 条、P C T 規則43、44〕

出願人又は代理人 の書類記号	SK01PCT13	今後の手続きについては、国際調査報告の送付通知様式(P C T / I S A / 2 2 0) 及び下記 5 を参照すること。	
国際出願番号 P C T / J P 0 1 / 0 1 1 2 7	国際出願日 (日.月.年)	1 6 . 0 2 . 0 1	優先日 (日.月.年) 2 8 . 0 2 . 0 0
出願人 (氏名又は名称) ソニー株式会社			

国際調査機関が作成したこの国際調査報告を法施行規則第41条 (P C T 1 8 条) の規定に従い出願人に送付する。
この写しは国際事務局にも送付される。

この国際調査報告は、全部で 2 ページである。

☐ この調査報告に引用された先行技術文献の写しも添付されている。

1. 国際調査報告の基礎

a. 言語は、下記に示す場合を除くほか、この国際出願がされたものに基づき国際調査を行った。

☐ この国際調査機関に提出された国際出願の翻訳文に基づき国際調査を行った。

b. この国際出願は、ヌクレオチド又はアミノ酸配列を含んでおり、次の配列表に基づき国際調査を行った。

☐ この国際出願に含まれる書面による配列表

☐ この国際出願と共に提出されたフレキシブルディスクによる配列表

☐ 出願後に、この国際調査機関に提出された書面による配列表

☐ 出願後に、この国際調査機関に提出されたフレキシブルディスクによる配列表

☐ 出願後に提出した書面による配列表が出願時における国際出願の開示の範囲を超える事項を含まない旨の陳述書の提出があった。

☐ 書面による配列表に記載した配列とフレキシブルディスクによる配列表に記載した配列が同一である旨の陳述書の提出があった。

2. ☐ 請求の範囲の一部の調査ができない (第 I 欄参照)。

3. ☐ 発明の単一性が欠如している (第 II 欄参照)。

4. 発明の名称は ☒ 出願人が提出したものを承認する。

☐ 次に示すように国際調査機関が作成した。

5. 要約は ☒ 出願人が提出したものを承認する。

☐ 第 III 欄に示されているように、法施行規則第47条 (P C T 規則38.2(b)) の規定により国際調査機関が作成した。出願人は、この国際調査報告の発送の日から 1 カ月以内にこの国際調査機関に意見を提出することができる。

6. 要約書とともに公表される図は、
第 2 図とする。 ☒ 出願人が示したとおりである。

☐ なし

☐ 出願人は図を示さなかった。

☐ 本図は発明の特徴を一層よく表している。

THIS PAGE BLANK (USPTO)

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int Cl' G10L15/28, 15/18

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int Cl' G10L15/00~17/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922~1996年
 日本国公開実用新案公報 1971~2001年
 日本国登録実用新案公報 1994~2001年
 日本国実用新案登録公報 1996~2001年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICST科学技術文献ファイル (JOIS)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	IEEE Transactions on Speech and Audio Processing, Vol.1, No.1, January 1993, L.R.Bahl et al, "A Fast Approximate Acoustic Match for Large Vocabulary Speech Recognition", p.59-67, ISSN1063-6676	1~9
A	J P, 57-86899, A (三菱電機株式会社), 31.5月.1982(31.05.82), 全文全図, (ファミリーなし)	1~9
A	J P, 58-52696, A (大日本印刷株式会社), 28.3月.1983(28.03.83), 全文全図, (ファミリーなし)	1~9

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」 口頭による開示、使用、展示等に言及する文献
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」 同一パテントファミリー文献

国際調査を完了した日

29.03.01

国際調査報告の発送日

10.04.01

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

松尾 淳

印

5C 8842

電話番号 03-3581-1101 内線 3540

THIS PAGE BLANK (USPTO)

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2001年9月7日 (07.09.2001)

PCT

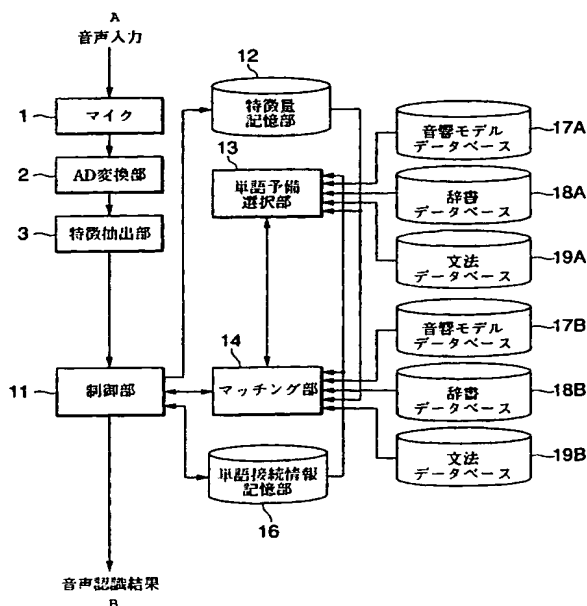
(10) 国際公開番号
WO 01/65541 A1

- (51) 国際特許分類⁷: G10L 15/28, 15/18
(21) 国際出願番号: PCT/JP01/01127
(22) 国際出願日: 2001年2月16日 (16.02.2001)
(25) 国際出願の言語: 日本語
(26) 国際公開の言語: 日本語
(30) 優先権データ:
特願2000-51466 2000年2月28日 (28.02.2000) JP
(71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 Tokyo (JP).
(72) 発明者; および
(75) 発明者/出願人 (米国についてののみ): 浅野康治
(74) 代理人: 小池 晃, 外 (KOIKE, Akira et al.) ; 〒105-0001 東京都港区虎ノ門二丁目6番4号 第11森ビル Tokyo (JP).
(81) 指定国 (国内): CN, JP, US.
(84) 指定国 (広域): ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).
添付公開書類:
— 国際調査報告書

[続葉有]

(54) Title: SPEECH RECOGNITION DEVICE AND SPEECH RECOGNITION METHOD, AND RECORDING MEDIUM

(54) 発明の名称: 音声認識装置及び音声認識方法、並びに記録媒体



(57) Abstract: A speech recognition device for improving a speech recognition accuracy while restricting an increase in resource, wherein a probable word as the result of speech recognition is selected based on acoustic scores and linguistic scores, and, in addition, a word is selected based on criteria other than acoustic scores such as "small number of phonemes", "being specified part of speech", "included in the past speech recognition results", and "linguistic score higher than a preset value", and the selected word is subjected to matching.

- A...VOICE INPUT
1...MICROPHONE
2...A/D CONVERTER
3...FEATURE EXTRACTING UNIT
11...CONTROL UNIT
B...SPEECH RECOGNITION RESULT
12...FEATURE QUANTITY STORING UNIT
13...WORD PRE-SELECTING UNIT
14...MATCHING UNIT
16...WORD CONNECTION INFORMATION STORING UNIT
17A...ACOUSTIC MODEL DATABASE
18A...DICTIONARY DATABASE
19A...GRAMMAR DATABASE
17B...ACOUSTIC MODEL DATABASE
18B...DICTIONARY DATABASE
19B...GRAMMAR DATABASE

[続葉有]



2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(57) 要約:

本発明は、リソースの増加を抑えて、音声認識精度を向上させる音声認識装置であり、音響スコア及び言語スコアに基づいて、音声認識結果として確からしい単語が選択される他、例えば、「音韻数が短い」、「所定の品詞である」、「過去の音声認識結果に含まれる」、「言語スコアが所定値以上である」等といった、音響スコアとは異なる尺度に基づいて単語が選択され、その選択された単語を対象にマッチング処理が行われる。

明細書

音声認識装置及び音声認識方法、並びに記録媒体

技術分野

本発明は、音声認識装置及び音声認識方法、並びに記録媒体に関し、特に、例えば、音響的な特徴量が不安定な単語を含む音声であっても、少ないリソースで精度の良い音声認識を行うことができるようにする音声認識装置及び音声認識方法、並びに記録媒体に関する。

背景技術

図 1 は、従来の音声認識装置の一例を示している。

ユーザが発した音声は、マイクロホン 1 に入力される。マイクロホン 1 は、入力された音声を、電気信号としての音声信号に変換する。この音声信号は、A/D (Analog/Digital) 変換部 2 に供給される。A/D 変換部 2 は、マイクロホン 1 から出力されるアナログ信号である音声信号をサンプリング、量子化し、ディジタル信号である音声データに変換する。この音声データは、特徴抽出部 3 に供給される。

特徴抽出部 3 は、A/D 変換部 2 からの音声データについて、適当なフレーム毎に音響処理を施し、これにより、例えば、M F C C (Mel Frequency Cepstrum Coefficient) 等の特徴量を抽出し、マッチング部 4 に供給する。なお、特徴抽出部 3 では、その他、例えば、スペクトルや、線形予測係数、ケプストラム係数、線スペクトル対等の特徴量を抽出することが可能である。

マッチング部 4 は、特徴抽出部 3 からの特徴量を用いて、音響モデルデータベース 5、辞書データベース 6 及び文法データベース 7 を必要に応じて参照しながら、マイクロホン 1 に入力された音声（入力音声）を、例えば、連続分布 HMM 法等に基づいて音声認識する。

即ち、音響モデルデータベース5は、音声認識する音声の言語における個々の音素や音節などの音響的な特徴を表す音響モデルを記憶している。ここでは、連続分布HMM法に基づいて音声認識を行うので、音響モデルとしては、HMM (Hidden Markov Model)が用いられる。辞書データベース6は、認識対象の各単語（語彙）について、その発音に関する情報（音韻情報）が記述された単語辞書を記憶している。文法データベース7は、辞書データベース6の単語辞書に登録されている各単語が、どのように連鎖、すなわちつながるかを記述した文法規則（言語モデル）を記憶している。ここで、文法規則としては、例えば、文脈自由文法（CFG）や、統計的な単語連鎖確率（N-gram）などに基づく規則を用いることができる。

マッチング部4は、辞書データベース6の単語辞書を参照することにより、音響モデルデータベース5に記憶されている音響モデルを接続することで、単語の音響モデル（単語モデル）を構成する。さらに、マッチング部4は、幾つかの単語モデルを、文法データベース7に記憶された文法規則を参照することにより接続し、そのようにして接続された単語モデルを用いて、特徴量に基づき、連続分布HMM法によって、マイクロホン1に入力された音声を認識する。即ち、マッチング部4は、特徴抽出部3が出力する時系列の特徴量が観測されるスコア（尤度）が最も高い単語モデルの系列を検出し、その単語モデルの系列に対応する単語列を音声の認識結果として出力する。

具体的には、マッチング部4は、接続された単語モデルに対応する単語列について、各特徴量の出現確率を累積し、その累積値をスコアとして、そのスコアを最も高くする単語列を音声認識結果として出力する。

スコア計算は、一般に、音響モデルデータベース5に記憶された音響モデルによって与えられる音響的なスコア（以下、適宜、音響スコアという）と、文法データベース7に記憶された文法規則によって与えられる言語的なスコア（以下、適宜、言語スコアという）とを総合評価することで行われる。

即ち、音響スコアは、例えば、HMM法による場合には、単語モデルを構成する音響モデルから、特徴抽出部3が出力する特徴量の系列が観測される確率（出現する確率）に基づいて、単語ごとに計算される。また、言語スコアは、例えば、

バイグラムによる場合には、注目している単語と、その単語の直前の単語とが連鎖（接続）する確率に基づいて求められる。そして、各単語についての音響スコアと言語スコアとを総合評価して得られる最終的なスコア（以下、適宜、最終スコアという）に基づいて音声認識結果が確定される。

具体的には、あるN個の単語からなる単語列におけるk番目の単語を w_k として、その単語 w_k の音響スコアを $A(w_k)$ と、言語スコアを $L(w_k)$ と、それぞれ表すとき、その単語列の最終スコアSは、例えば、式1にしたがって計算される。

$$S = \Sigma (A(w_k) + C_k \times L(w_k)) \cdots (1)$$

但し、 Σ は、kを1からNに変えてのサメーションをとることを表す。また、 C_k は、単語 w_k の言語スコア $L(w_k)$ にかけける重みを表す。

マッチング部4では、例えば、式1に示す最終スコアを最も大きくするNと、単語列 w_1, w_2, \dots, w_N を求めるマッチング処理が行われ、その単語列 w_1, w_2, \dots, w_N が、音声認識結果として出力される。

以上のような処理が行われることにより、図1の音声認識装置では、例えば、ユーザが、「ニューヨークに行きたいです」と発話した場合には、「ニューヨーク」、「に」、「行きたい」、「です」といった各単語に、音響スコア及び言語スコアが与えられ、それらを総合評価して得られる最終スコアが最も大きいときと、単語列「ニューヨーク」、「に」、「行きたい」、「です」が音声認識結果として出力される。

ところで、上述の場合において、辞書データベース6の単語辞書に、「ニューヨーク」、「に」、「行きたい」及び「です」の5単語が登録されているとすると、これらの5単語を用いて構成し得る5単語の並びは、 5^5 通り存在する。従って、単純には、マッチング部4では、この 5^5 通りの単語列を評価し、その中から、ユーザの発話に最も適合するもの、すなわち最終スコアを最も大きくするものを決定しなければならない。単語辞書に登録する単語数が増えれば、その単語数分の単語の並びの数は、単語数の単語数乗通りになるから、評価の対象としなければならない単語列は膨大な数となる。

さらに、一般には、発話中に含まれる単語の数は未知であるから、5単語の並びからなる単語列だけでなく、1単語、2単語、 \dots からなる単語列も、評価

の対象とする必要がある。従って、評価すべき単語列の数は、さらに膨大なものとなるから、そのような膨大な単語列の中から、音声認識結果として最も確からしいものを、計算量及び使用するメモリ容量の観点から効率的に決定することは、非常に重要な問題である。

計算量及びメモリ容量の効率化を図る方法としては、例えば、音響スコアを求める過程において、その途中で得られる音響スコアが所定の閾値以下となった場合に、そのスコア計算を打ち切るという音響的な枝刈り手法や、言語スコアに基づいて、スコア計算の対象とする単語を絞り込む言語的な枝刈り手法がある。

これらの枝刈り手法によれば、スコア計算の対象が、所定の判断基準（例えば、上述したような計算途中の音響スコアや、単語に与えられる言語スコア）に基づいて絞り込まれることで、計算量の削減を図ることができる。しかしながら、その反面、絞り込みを強くすると、即ち、判断基準を厳しくすると、本来、音声認識結果として正しいものまでも枝刈りされてしまい、誤認識が生じることになる。従って、枝刈り手法による場合には、音声認識結果として正しいものが枝刈りされないように、ある程度のマージンをもたせた絞り込みを行う必要があり、このため、計算量を大きく削減することは困難である。

また、音響スコアを求める場合に、スコア計算の対象となっているすべての単語について独立に行うと、その計算量が大きくなることから、複数の単語について音響スコアの計算の一部を共通化（共有化）する方法が提案されている。この共通化の方法としては、単語辞書の単語のうち、その先頭の音韻が同一のものについて、その先頭の音韻から、同一になっている音韻までは、音響モデルを共通に用い、それ以後の異なる音韻には、音響モデルを個々に用いることにより、全体として1つの木構造のネットワークを構成し、これを用いて、音響スコアを求める方法がある。具体的には、例えば、いま、単語「秋田」と「曙」を考え、「秋田」の音韻情報が「akita」であり、「曙」の音韻情報が「akebono」であるとする、単語「秋田」と「曙」の音響スコアは、それぞれの先頭から2番目までの音韻a,kについては兼用で計算される。そして、単語「秋田」の残りの音韻k,i,t,a、及び単語「曙」の残りの音韻e,b,o,n,oについては、それぞれ独立に音響スコアが計算される。

従って、この方法によれば、音響スコアの計算量を大幅に低減することができる。

この方法では、共通化されている部分である音響スコアが兼用で計算される部分において、その音響スコアの計算の対象となっている単語を決定することができない。即ち、上述の単語「秋田」と「曙」の例でいえば、それぞれの先頭から2番目までの音韻a,kについて音響スコアが計算されている場合は、その音響スコアが計算されている単語が、「秋田」であるのか、又は「曙」であるのかを同定することができない。

この場合、「秋田」については、その3番目の音韻iについて音響スコアの計算が開始されたときに、その計算対象が「秋田」であることを同定することができ、「曙」についても、その3番目の音韻eについての音響スコアの計算が開始されたときに、その計算対象が「曙」であることを同定することができる。

従って、音響スコアの計算の一部を共通化してしまうと、単語の音響スコアの計算の開始時に、その単語を同定することができないため、その単語について、言語スコアを考慮することができない。その結果、単語の音響スコアの開始前に、上述したような言語的な枝刈り手法を用いることが困難となり、無駄な計算が行われることがある。

さらに、音響スコアの計算の一部を共通化する場合、単語辞書のすべての単語を対象として、上述したような木構造のネットワークが構成されるから、これを保持するための大きなメモリ容量が必要となる。

また、計算量及びメモリ容量の効率化を図る方法としては、音響スコアを計算する場合に、単語辞書のすべての単語を対象とするのではなく、その音響スコアの計算の対象とする単語を予備的に選択（予備選択）し、その予備選択された単語についてだけ、音響スコアを計算する方法がある。

ここで、予備選択の方法は、例えば、L. R. Bahl, S. V. De Gennaro, P. S. Gopalakrishnan and R. L. Mercer, "A Fast Approximate Acoustic Match for Large Vocabulary Speech Recognition", IEEE Trans. Speech and Audio Proc., vol. 1, pp.59-67, 1993等に記載されている。

予備選択は、一般に、それほど精度の低い、簡易的な音響モデルや文法規

則を用いて行われる。即ち、予備選択は、単語辞書の単語すべてを対象として行われるため、精度の高い音響モデルや文法規則を用いて予備選択を行うと、リアルタイム性等を維持するのに、計算量やメモリ容量といったリソースが多く必要となる。そこで、予備選択は、簡易的な音響モデルや文法規則を用いることで、大語彙を対象とした場合でも、比較的少ないリソースで、高速に行うことが可能となっている。

予備選択を行う音声認識装置では、予備選択された単語についてだけマッチング処理を行えば良いので、マッチング処理は、精度の高い音響モデルや文法規則を用いても、少ないリソースで、高速に行うことができる。従って、予備選択を行う音声認識装置は、大語彙を対象として音声認識を行う場合に、特に有用である。

ところで、予備選択は、ある単語について、特徴量の系列（特徴量系列）を用いてのマッチング処理が終了し、とりあえず確からしい終点が求められた後に、その終点を始点として、その始点に対応する時刻以後の特徴量系列を用いて行われる。即ち、予備選択は、連続発話された音声に含まれる単語どうしの境界（単語境界）が、最終的に確定していない時点で行われる。

従って、予備選択に用いられる特徴量系列の始点や終点が、対応する単語の始点や終点からずれている場合には、その単語の直前の単語や直後の単語の音韻の特徴量を含む特徴量系列や、対応する単語の最初や最後の部分の特徴量が欠けた特徴量系列、即ち、いわば音響的に安定していない特徴量系列を用いて、予備選択が行われることになる。

このため、簡易的な音響モデルを用いる予備選択では、発話中に含まれる単語が選択されないことが起こり得る。特に、例えば、日本語の助詞や助動詞、英語の冠詞や前置詞などの音韻数が短い単語については、そのような選択漏れが生じる可能性が高い。

そして、予備選択において、正しい単語が選択されない場合には、その単語についてマッチング処理が行われないから、音声認識結果は誤ったものとなる。

そこで、予備選択において、単語を選択するときの音響的又は言語的な判断基準を緩くして、選択される単語の数を多くする方法や、精度の高い音響モデル及

び文法規則を用いる方法がある。

しかしながら、予備選択において、単語を選択するときの音響的又は言語的な判断基準を緩くすると、音声認識結果としてそれほど可能性の低い単語の多くが、マッチング処理の対象となり、予備選択に比較して1単語あたりの負荷の重いマッチング処理に要するリソースが大きく増大する。

また、予備選択において、精度の高い音響モデル及び文法規則を用いる場合には、予備選択に要するリソースが大きく増大する。

発明の開示

本発明は、このような状況に鑑みてなされたものであり、必要なリソースの増加を極力抑えながら、音声認識精度を向上させることができるようにするものである。

本発明に係る音声認識装置は、音声認識の対象とする単語群から、1以上の第1の単語を、特徴量を用いて計算される第1の尺度に基づいて選択するとともに、1以上の第2の単語を、第1の尺度とは異なる第2の尺度に基づいて選択する選択手段と、選択手段において選択された第1及び第2の単語について、スコアを計算するスコア計算手段とを備える。

選択手段には、音韻数を、第2の尺度として、音韻数が所定の条件を満たす単語を、第2の単語として選択させることができる。また、選択手段には、品詞を、第2の尺度として、品詞が所定の条件を満たす単語を、第2の単語として選択させることができる。さらに、選択手段には、言語的な尤度を、第2の尺度として、言語的な尤度が高い単語を、第2の単語として選択させることができる。また、本発明の音声認識装置には、音声認識結果を記憶する記憶手段をさらに備えることができ、この場合、選択手段には、記憶手段に記憶されていることを、第2の尺度として、記憶手段に記憶されている音声認識結果に含まれる単語を、第2の単語として選択させることができる。

また、本発明に係る音声認識装置は、音声認識結果を修正するための入力を与える入力手段をさらに備えることができる。この場合、記憶手段は、入力手段か

らの入力にしたがって修正された音声認識結果を記憶する。また、選択手段は、音声の特徴量を用いて、スコアを計算させ、そのスコアに基づいて、第1の単語を選択する。

本発明に係る音声認識方法は、音声認識の対象とする単語群から、1以上の第1の単語を、特徴量を用いて計算される第1の尺度に基づいて選択するとともに、1以上の第2の単語を、第1の尺度とは異なる第2の尺度に基づいて選択する選択ステップと、選択ステップにおいて選択された第1及び第2の単語について、スコアを計算するスコア計算ステップとを備える。

本発明に係る記録媒体は、音声認識の対象とする単語群から、1以上の第1の単語を、特徴量を用いて計算される第1の尺度に基づいて選択するとともに、1以上の第2の単語を、第1の尺度とは異なる第2の尺度に基づいて選択する選択ステップと、選択ステップにおいて選択された第1及び第2の単語について、スコアを計算するスコア計算ステップとを備えるプログラムが記録されている。

本発明に係る音声認識装置及び音声認識方法、並びに記録媒体は、音声認識の対象とする単語群から、1以上の第1の単語が、特徴量を用いて計算される第1の尺度に基づいて選択されるとともに、1以上の第2の単語が、第1の尺度とは異なる第2の尺度に基づいて選択され、その選択された第1及び第2の単語について、スコアが計算される。

図面の簡単な説明

図1は、従来の音声認識装置の一例を示すブロック図である。

図2は、本発明に係る音声認識装置の一例を示すブロック図である。

図3は、単語接続情報を説明するための図である。

図4は、音声認識装置の処理を説明するためのフローチャートである。

図5は、本発明に係る音声認識装置の他の例を示すブロック図である。

図6は、音声認識装置の処理を説明するためのフローチャートである。

図7は、本発明を適用したコンピュータの一例を示すブロック図である。

発明を実施するための最良の形態

図 2 は、本発明に係る音声認識装置の一例を示している。なお、図中、図 1 に示す装置と対応する部分については、同一の符号を付し、以下では、その詳細な説明は適宜省略する。

特徴量抽出部 3 が出力するユーザが発した音声の特徴量の系列は、フレーム単位で制御部 11 に供給され、制御部 11 は、特徴量抽出部 3 からの特徴量を特徴量記憶部 12 に供給する。

制御部 11 は、単語接続情報記憶部 16 に記憶された単語接続情報を参照してマッチング部 14 を制御する。さらに、制御部 11 は、マッチング部 14 が前述した図 1 のマッチング部 4 と同様のマッチング処理を行うことにより得られるマッチング処理結果としての音響スコアや言語スコア等に基づいて、単語接続情報を生成し、その単語接続情報によって、単語接続情報記憶部 16 の記憶内容を更新する。また、制御部 11 は、単語接続情報記憶部 16 に記憶された単語接続情報に基づいて最終的な音声認識結果を確定して出力する。

特徴量記憶部 12 は、制御部 11 から供給される特徴量の系列を、例えば、ユーザの音声の認識結果が得られるまで記憶する。なお、制御部 11 は、音声区間の開始時刻を基準（例えば 0）とする、特徴抽出部 3 が出力する特徴量が得られた時刻（以下、適宜、抽出時刻という）を、その特徴量とともに特徴量記憶部 12 に供給するようになっており、特徴量記憶部 12 は、特徴量をその抽出時刻とともに記憶する。特徴量記憶部 12 に記憶された特徴量及びその抽出時刻は、単語予備選択部 13 及びマッチング部 14 において必要に応じて参照することができるようになっている。

単語予備選択部 13 は、マッチング部 14 からの要求に応じ、単語接続情報記憶部 16、音響モデルデータベース 17 A、辞書データベース 18 A 及び文法データベース 19 A を必要に応じて参照しながら、マッチング部 14 でマッチング処理の対象とする 1 以上の単語を選択する単語予備選択処理を、特徴量記憶部 12 に記憶された特徴量を用いて行う。

マッチング部 14 は、制御部 11 からの制御に基づき、単語接続情報記憶部 1

6、音響モデルデータベース 17 B、辞書データベース 18 B 及び文法データベース 19 B を必要に応じて参照しながら、単語予備選択部 13 からの単語予備選択処理の結果得られる単語を対象としたマッチング処理を、特徴量記憶部 12 に記憶された特徴量を用いて行い、そのマッチング処理の結果を制御部 11 に供給する。

単語接続情報記憶部 16 は、制御部 11 から供給される単語接続情報をユーザの音声の認識結果が得られるまで記憶する。

ここで、単語接続情報は、最終的な音声認識結果の候補となる単語列を構成する単語どうしの接続（連鎖又は連接）関係を表すもので、各単語の音響スコア及び言語スコア並びに各単語に対応する発話の開始時刻及び終了時刻も含んでいる。

即ち、図 3 は、単語接続情報記憶部 16 に記憶される単語接続情報を、グラフ構造を用いて示している。

図 3 に示す例において、単語接続情報としてのグラフ構造は、単語を表すアーク（図 3 において、○印どうしを結ぶ線分で示す部分）と、単語どうしの境界を表すノード（図 3 において○印で示す部分）とから構成されている。

ノードは、時刻情報を有しており、この時刻情報は、そのノードに対応する特徴量の抽出時刻を表す。上述したように、抽出時刻は、音声区間の開始時刻を 0 とする、特徴抽出部 3 が出力する特徴量が得られた時刻であるから、図 3 において、音声区間の開始、即ち、最初の単語の先頭に対応するノード Node₁ が有する時刻情報は 0 となる。ノードは、アークの始端及び終端となるが、始端のノード（始端ノード）、又は終端のノード（終端ノード）が有する時刻情報は、それぞれ、そのノードに対応する単語の発話の開始時刻、又は終了時刻となる。

なお、図 3 では、左から右方向が、時間の経過を表しており、従って、あるアークの左右にあるノードのうち、左側のノードが始端ノードとなり、右側のノードが終端ノードとなる。

アークは、そのアークに対応する単語の音響スコア及び言語スコアを有しており、このアークが、終端ノードとなっているノードを始端ノードとして、順次接続されていくことにより、音声認識結果の候補となる単語の系列が構成されていく。

即ち、制御部 11 においては、まず最初に、音声区間の開始を表すノード Node₁ に対して、音声認識結果として確からしい単語に対応するアークが接続される。図 3 の実施の形態では、「今日」に対応するアーク Arc₁、「いい」に対応するアーク Arc₆、及び「天気」に対応する Arc₁₁が接続されている。なお、音声認識結果として確からしい単語かどうかは、マッチング部 14 において求められる音響スコア及び言語スコアに基づいて決定される。

そして、以下、同様にして、「今日」に対応するアーク Arc₁の終端である終端ノード Node₂、「いい」に対応するアーク Arc₆の終端である終端ノード Node₇、「天気」に対応する Arc₁₁の終端である終端ノード Node₁₂それぞれに対して、同様に、確からしい単語に対応するアークが接続されていく。

以上のようにしてアークが接続されていくことで、音声区間の開始を始点として、左から右方向に、アークとノードで構成される 1 以上のバスが構成されて行くが、例えば、そのバスのすべてが、音声区間の最後（図 3 では、時刻 T）に到達すると、制御部 11 において、音声区間の開始から最後までに形成された各バスについて、そのバスを構成するアークが有している音響スコア及び言語スコアが累積され、最終スコアが求められる。そして、例えば、その最終スコアが最も高いバスを構成するアークに対応する単語列が、音声認識結果として確定されて出力される。

具体的には、図 3 において、例えば、ノード Node 1 から、「今日」に対応するアーク Arc₁、ノード Node₂、「は」に対応するアーク Arc₂、ノード Node₃、「いい」に対応するアーク Arc₃、ノード Node₄、「天気」に対応するアーク Arc₄、ノード Node₅、「ですね」に対応するアーク Arc₅、及びノード Node₆で構成されるバスについて、最も高い最終スコアが得られた場合には、単語列「今日」、「は」、「いい」、「天気」、「ですね」が、音声認識結果として出力されることになる。

なお、上述の場合には、音声区間内にあるノードについて、必ずアークを接続して、音声区間の開始から最後にまで延びるバスを構成するようにしたが、このようなバスを構成する過程において、それまでに構成されたバスについてのスコアから、音声認識結果として不適當であることが明らかであるバスに関しては、その時点で、バスの構成を打ち切る（その後に、アークを接続しない）ようにす

ることが可能である。

上述のようなバスの構成ルールに従えば、1つのアークの終端が、次に接続される1以上のアークの始端ノードなり、基本的には、枝葉が広がるように、バスが構成されて行くが、例外的に、1つのアークの終端が、他のアークの終端に一致する場合、つまり、あるアークの終端ノードと、他のアークの終端ノードとが同一のノードに共通化される場合がある。

文法規則として、例えば、バイグラムを用いた場合には、別のノードから延びる2つのアークが、同一の単語に対応するものであり、さらに、その単語の発話の終了時刻も同一であるときには、その2つのアークの終端は一致する。

図3において、ノードNode₇を始端として延びるアークArc₇、及びノードNode₁₃を始端として延びるアークArc₁₃は、いずれも「天気」に対応するものであり、その発話の終了時刻も同一であるため、その終端ノードは、同一のノードNode₈に共通化されている。

なお、ノードの共通化は行わないようにすることも可能であるが、メモリ容量の効率化の観点からは、行うのが好ましい。

図3では、文法規則としてバイグラムを用いているが、その他、例えば、トライグラム等を用いる場合も、ノードの共通化は可能である。

さらに、単語接続情報記憶部16に記憶されている単語接続情報は、単語予備選択部13及びマッチング部14において、必要に応じて参照することができるようになっている。

図2に戻り、音響モデルデータベース17A及び17Bは、基本的には、図1の音響モデルデータベース5において説明したような音響モデルを記憶している。

但し、音響モデルデータベース17Bは、音響モデルデータベース17Aよりも精度の高い処理が可能な高精度の音響モデルを記憶している。即ち、音響モデルデータベース17Aにおいて、各音素や音節について、例えば、前後のコンテキストに依存しない1パターンの音響モデルだけが記憶されているとすると、音響モデルデータベース17Bには、各音素や音節について、例えば、前後のコンテキストに依存しない音響モデルの他、単語間にまたがるコンテキストに依存する音響モデル、つまり、クロスワードモデルや、単語内のコンテキストに依存す

る音響モデルも記憶されている。

辞書データベース 18 A 及び 18 B は、基本的には、図 1 の辞書データベース 6 において説明したような単語辞書を記憶している。

即ち、辞書データベース 18 A 及び 18 B の単語辞書には、同一セットの単語が登録されている。但し、辞書データベース 18 B の単語辞書は、辞書データベース 18 A の単語辞書よりも精度の高い処理が可能な高精度の音韻情報を記憶している。即ち、辞書データベース 18 A の単語辞書には、例えば、各単語に対して、1 通りの音韻情報（読み）だけ登録されているとすると、辞書データベース 18 B の単語辞書には、例えば、各単語に対して、複数通りの音韻情報が登録されている。

具体的には、例えば、単語「お早う」に対して、辞書データベース 18 A の単語辞書には、1 通りの音韻情報「おはよう」だけが、辞書データベース 18 B の単語辞書には、「おはよう」の他、「おはよー」や「おはよ」が、それぞれ音韻情報として登録されている。

文法データベース 19 A 及び 19 B は、基本的には、図 1 の文法データベース 7 において説明したような文法規則を記憶している。

但し、文法データベース 19 B は、文法データベース 19 A よりも精度の高い処理が可能な高精度の文法規則を記憶している。即ち、文法データベース 19 A が、例えば、ユニグラム（単語の生起確率）に基づく文法規則を記憶しているとすると、文法データベース 19 B は、例えば、バイグラム（直前の単語との関係を考慮した単語の生起確率）や、トライグラム（直前の単語及びそのさらに 1 つ前の単語との関係を考慮した単語の生起確率）、文脈自由文法等に基づく文法規則を記憶している。

以上のように、音響モデルデータベース 17 A には、各音素や音節について、1 パターンの音響モデルが、音響モデルデータベース 17 B には、各音素や音節について、複数パターンの音響モデルが、それぞれ記憶されている。また、辞書データベース 18 A には、各単語について、1 通りの音韻情報が、辞書データベース 18 B には、各単語について、複数通りの音韻情報が、それぞれ記憶されている。そして、文法データベース 19 A には、簡易な文法規則が、文法データベ

ース 19 B には、精度の高い文法規則が、それぞれ記憶されている。

これにより、音響モデルデータベース 17 A、辞書データベース 18 A 及び文法データベース 19 A を参照する単語予備選択部 13 では、それほど精度は高くないが、多くの単語を対象として、迅速に、音響スコアや言語スコアを求めることができる。また、音響モデルデータベース 17 B、辞書データベース 18 B、及び文法データベース 19 B を参照するマッチング部 14 では、ある程度の数の単語を対象として、迅速に、精度の高い音響スコアや言語スコアを求めることができる。

ここでは、音響モデルデータベース 17 A 及び 17 B それぞれに記憶させる音響モデルの精度について優劣を設けるようにしたが、音響モデルデータベース 17 A 及び 17 B には、いずれにも同一の音響モデルを記憶させることができ、この場合、音響モデルデータベース 17 A 及び 17 B は、1 つの音響モデルデータベースに共通化することができる。同様に、辞書データベース 18 A 及び 18 B の単語辞書それぞれの記憶内容や、文法データベース 19 A 及び 19 B それぞれの文法規則も同一にすることができる。

次に、図 4 のフローチャートを参照して、図 2 の音声認識装置による音声認識処理について説明する。

ユーザが発話を行うと、その発話としての音声は、マイクロホン 1 及び A/D 変換部 2 を介することにより、デジタルの音声データとされ、特徴抽出部 3 に供給される。特徴抽出部 3 は、そこに供給される音声データから、音声の特徴量を、フレームごとに順次抽出し、制御部 11 に供給する。

制御部 11 は、何らかの手法で音声区間を認識するようになっており、音声区間においては、特徴抽出部 3 から供給される特徴量の系列を、各特徴量の抽出時刻と対応付けて、特徴量記憶部 12 に供給して記憶させる。

さらに、制御部 11 は、音声区間の開始後、ステップ S1 において、音声区間の開始を表すノード（以下、適宜、初期ノードという）を生成し、単語接続情報記憶部 16 に供給して記憶させる。即ち、制御部 11 は、ステップ S1 において、図 3 におけるノード Node₁ を、単語接続情報記憶部 16 に記憶させる。

そして、ステップ S2 に進み、制御部 11 は、単語接続情報記憶部 16 の単語

接続情報を参照することで、途中ノードが存在するかどうかを判定する。

即ち、上述したように、図3に示した単語接続情報においては、終端ノードに、アークが接続されていくことにより、音声区間の開始から最後にまで延びるパスが形成されて行くが、ステップS2では、終端ノードのうち、まだアークが接続されておらず、かつ、音声区間の最後にまで到達していないものが、途中ノード（例えば、図3におけるノードNode₈や、Node₁₀、Node₁₁）として検索され、そのような途中ノードが存在するかどうか判定される。

上述したように、音声区間は何らかの手法で認識され、さらに、終端ノードに対応する時刻は、その終端ノードが有する時刻情報を参照することで認識することができるから、アークが接続されていない終端ノードが、音声区間の最後に到達していない途中ノードであるかどうかは、音声区間の最後の時刻と、終端ノードが有する時刻情報とを比較することで判定することができる。

ステップS2において、途中ノードが存在すると判定された場合、ステップS3に進み、制御部11は、情報接続情報の中に存在する途中ノードのうちの1つを、それに接続するアークとしての単語を決定するノード（以下、適宜、注目ノードという）として選択する。

即ち、制御部11は、情報接続情報の中に1つの途中ノードしか存在しない場合には、その途中ノードを注目ノードとして選択する。また、制御部11は、情報接続情報の中に複数の途中ノードが存在する場合には、複数の途中ノードのうちの1つを注目ノードとして選択する。具体的には、制御部11は、例えば、複数の途中ノードそれぞれが有する時刻情報を参照し、その時刻情報が表す時刻が最も古いもの（音声区間の開始側のもの）、又は最も新しいもの（音声区間の終わり側のもの）を、注目ノードとして選択する。また、制御部11は、例えば、初期ノードから、複数の途中ノードそれぞれに至るまでのパスを構成するアークが有する音響スコア及び言語スコアを累積し、その累積値（以下、適宜、部分累積スコアという）が最も大きくなるパス、又は小さくなるパスの終端になっている途中ノードを、注目ノードとして選択する。

その後、制御部11は、注目ノードが有する時刻情報を開始時刻としてマッチング処理を行う旨の指令（以下、適宜、マッチング処理指令という）を、マッチ

ング部 1 4 に出力する。

マッチング部 1 4 は、制御部 1 1 からマッチング処理指令を受信すると、注目ノード、及びそれが有する時刻情報を、単語予備選択部 1 3 に供給し、単語予備選択処理を要求して、ステップ S 4 に進む。

ステップ S 4 では、単語予備選択部 1 3 は、マッチング部 1 4 から、単語予備選択処理の要求を受信すると、注目ノードに接続されるアークとなる単語の候補を選択する単語予備選択処理を、辞書データベース 1 8 A の単語辞書に登録された単語を対象として行う。

即ち、単語予備選択部 1 3 は、言語スコア及び音響スコアを計算するのに用いる特徴量の系列の開始時刻を注目ノードが有する時刻情報から認識し、その開始時刻以降の必要な特徴量の系列を特徴量記憶部 1 2 から読み出す。さらに、単語予備選択部 1 3 は、辞書データベース 1 8 A の単語辞書に登録された各単語の単語モデルを音響モデルデータベース 1 7 A の音響モデルを接続することで構成し、その単語モデルに基づき、特徴量記憶部 1 2 から読み出した特徴量の系列を用いて音響スコアを計算する。

単語予備選択部 1 3 は、各単語モデルに対応する単語の言語スコアを文法データベース 1 9 A に記憶された文法規則に基づいて計算する。即ち、単語予備選択部 1 3 は、各単語の言語スコアを例えばユニグラムに基づいて求める。

単語予備選択部 1 3 においては、単語接続情報を参照することにより、各単語の音響スコアの計算をその単語の直前の単語（注目ノードが終端となっているアークに対応する単語）に依存するクロスワードモデルを用いて行うことが可能である。但し、クロスワードモデルを用いる場合には、用いない場合に比較して音響スコアの計算量は増大することになる。

また、単語予備選択部 1 3 においては、単語接続情報を参照することにより、各単語の言語スコアの計算を、その単語が、その直前の単語と連鎖する確率を規定するバイグラムに基づいて行うことが可能である。但し、バイグラム（さらには、トライグラム等）に基づいて言語スコアの計算を行う場合には、ユニグラムに基づいて行う場合に比較してその計算量は増大することになる。

単語予備選択部 1 3 は、以上のようにして、各単語について音響スコア及び言

語スコアを求めると、その音響スコア及び言語スコアを総合評価したスコア（以下、適宜、単語スコアという）（第1の尺度）を求め、その上位L個を、マッチング処理の対象とする単語としてマッチング部14に供給する。

さらに、単語予備選択部13は、ステップS4において、辞書データベース18Aに登録されている1以上の単語を、音響スコアが反映された単語スコアとは異なる尺度（第2の尺度）に基づいて選択してマッチング部14に供給する。

即ち、単語予備選択部13は、辞書データベース18Aに登録されている単語のうち、例えば、音素数又は音韻数が、所定値以下の短い単語及び所定の品詞の単語（例えば、英語における前置詞や冠詞、日本語における助詞や助動詞など）などの一般に発話時間が短い単語をその音響スコアに関係なく選択してマッチング部14に供給する。この場合、発話時間が短い単語は、必ず、マッチング処理の対象とされる。

単語予備選択部13において、上述のように、音響スコアに関係のない特定の尺度で選択される単語を、以下、適宜、特定単語という。

マッチング部14は、単語予備選択部13から、単語スコアに基づいて選択されたL個の単語と、単語スコアに関係ない所定の条件（尺度）に基づいて選択された特定単語を受信すると、ステップS5において、それらの単語を対象としてマッチング処理を行う。

即ち、マッチング部14は、言語スコア及び音響スコアを計算するのに用いる特徴量の系列の開始時刻を注目ノードが有する時刻情報から認識し、その開始時刻以降の必要な特徴量の系列を特徴量記憶部12から読み出す。さらに、マッチング部14は、辞書データベース18Bを参照することで、単語予備選択部13からの単語の音韻情報を認識し、その音韻情報に対応する音響モデルを音響モデルデータベース17Bから読み出して接続することで単語モデルを構成する。

マッチング部14は、上述のようにして構成した単語モデルに基づき、特徴量記憶部12から読み出した特徴量系列を用いて単語予備選択部13からの単語の音響スコアを計算する。マッチング部14は、単語接続情報を参照することにより、単語の音響スコアの計算をクロスワードモデルに基づいて行うようにすることが可能である。

さらに、マッチング部 14 は、文法データベース 19 B を参照することで、単語予備選択部 13 からの単語の言語スコアを計算する。即ち、マッチング部 14 は、例えば、単語接続情報を参照することにより、単語予備選択部 13 からの単語の直前の単語と、さらにその前の単語を認識し、トライグラムに基づく確率から単語予備選択部 13 からの単語の言語スコアを求める。

マッチング部 14 は、以上のようにして、単語予備選択部 13 からの L 個の単語と、特定単語のすべて（以下、適宜、これらをまとめて、選択単語という）についてその音響スコア及び言語スコアを求め、ステップ S 6 に進む。ステップ S 6 では、選択単語それぞれについてその音響スコア及び言語スコアを総合評価した単語スコアが求められ、その単語スコアに基づいて単語接続情報記憶部 16 に記憶された単語接続情報が更新される。

即ち、ステップ S 6 では、マッチング部 14 は、選択単語について単語スコアを求め、例えば、その単語スコアを所定の閾値と比較すること等によって、注目ノードに接続するアークとしての単語を選択単語の中から絞り込む。マッチング部 14 は、その絞り込みの結果残った単語を、その音響スコア、言語スコア及びその単語の終了時刻とともに、制御部 11 に供給する。

なお、単語の終了時刻は、音響スコアを計算するのに用いた特徴量の抽出時刻から認識される。また、ある単語について、その終了時刻としての蓋然性の高い抽出時刻が複数得られた場合には、その単語については、各終了時刻と、対応する音響スコア及び言語スコアとのセットが、制御部 11 に供給される。

制御部 11 は、上述のようにしてマッチング部 14 から供給される単語の音響スコア、言語スコア、及び終了時刻を受信すると、マッチング部 14 からの各単語について、単語接続情報記憶部 16 に記憶された単語接続情報（図 3）における注目ノードを始端ノードとして、アークを延ばし、そのアークを、終了時刻の位置に対応する終端ノードに接続する。さらに、制御部 11 は、各アークに対して、対応する単語、並びにその音響スコア及び言語スコアを付与するとともに、各アークの終端ノードに対して、対応する終了時刻を時刻情報として与える。そして、ステップ S 2 に戻り、以下、同様の処理が繰り返される。

以上のように、単語接続情報は、マッチング部 14 の処理結果に基づいて、逐

次更新されるので、単語予備選択部 1 3 及びマッチング部 1 4 は、常時、最新の単語接続情報を利用して処理を行うことが可能となる。

なお、制御部 1 1 は、単語接続情報を更新する際に、可能であれば、上述したような終端ノードの共通化を行う。

一方、ステップ S 2 において、途中ノードが存在しないと判定された場合、ステップ S 7 に進み、制御部 1 1 は、単語接続情報を参照することで、その単語接続情報として構成された各パスについて単語スコアを累積することで、最終スコアを求め、例えば、その最終スコアが最も大きいパスを構成するアークに対応する単語列をユーザの発話に対する音声認識結果として出力して処理を終了する。

以上のように、単語予備選択部 1 3 において、音響的な特徴量から計算される音響スコアを含む単語スコアに基づいて音声認識結果として確からしい単語を選択する他、例えば、日本語の助詞や助動詞、英語の冠詞や前置詞、その他の音韻数が短い音響的な特徴量が不安定な単語を音響的な特徴量から計算される音響スコアとは関係ない尺度に基づいて選択し、マッチング部 1 4 において、それらの単語をマッチング処理の対象とするようにしたので、音響的な特徴量が不安定な単語が単語予備選択部 1 3 で選択されないことによる音声認識精度の劣化を防止することができる。

この場合、単語予備選択部 1 3 において、単語を選択するときの枝刈り（絞り込み）の範囲を小さくしたり、精度の高い音響モデル及び文法規則を用いているわけではないので処理に必要なリソース（演算量やメモリ等）を極力低減することができる。

また、音響的な特徴量が不安定な音韻数の短い単語が、必ずマッチング処理の対象とされるため、単語予備選択部 1 3 において、音響スコアを含む単語スコアに基づいて選択される単語は、音響的な特徴量が比較的安定している音韻数の長い単語だけであっても良い。単語予備選択部 1 3 では、より簡易な音響モデルや文法規則を用いても、正しい単語の選択漏れが生じないこととなり、その結果、単語予備選択部 1 3 の処理に必要なリソースを低減しながら音声認識精度を向上させることができる。

さらに、単語予備選択部 1 3 において、音響スコアを含む単語スコアに基づい

て選択される単語が、音響的な特徴量が比較的安定している音韻数の長い単語だけとなる場合には、単語を選択するときの音響的又は言語的な判断基準としてより厳しいものを用い、音響スコアや言語スコアに基づいて選択される単語の数を少なくとも正しい単語の選択漏れが生じないこととなり、その結果、マッチング部 14 の処理に必要なリソースを低減しながら音声認識精度を向上させることができる。

上述の場合には、単語予備選択部 13 において、発話時間という尺度に基づいて、発話時間が短い単語をその音響スコアに関係なく選択するようにすることで、音響的な特徴量が不安定な単語が、単語予備選択部 13 で選択されないことによる音声認識精度の劣化を防止するようにしたが、その他、単語予備選択部 13 には、単語スコアに基づいて選択される単語とは別に、音響スコアとは関係がない尺度、即ち、音声の特徴量とは無関係の尺度である、例えば、言語スコアにのみ基づいて単語を選択させるようにすることができ、この場合もリソースを低減しながら音声認識精度を向上させることができる。

即ち、単語予備選択部 13 には、例えば、注目ノードに至るまでのパスの、注目ノード側の $N-1$ 単語の系列に接続する可能性の高い単語を、統計的な単語連鎖確率 ($N-gram$) に基づく文法規則から得られる言語スコアにのみ基づいて選択させることができる。この場合、特徴量から計算される音響スコアは考慮されずに単語が選択されることから、音響的な特徴量が不安定であることによって音響スコアが小さくなり、それに起因して単語スコアも小さくなってしまう単語が選択されないことによる音声認識精度の劣化を防止することができる。

また、言語スコアの計算は、音響スコアの計算に比較して計算量が非常に少なくて済むことから、リソースの増加を最小限に抑えることができる。

上述のように、統計的な単語連鎖確率に基づく文法規則から得られる言語スコアのみに基づいて行う単語の選択は、文法データベース 19A にそのような文法規則を記憶しておくことで容易に行うことが可能である。

次に、図 5 は、本発明に係る音声認識装置の他の例を示すブロック図であり、図 2 に示すものと共通する部分については、同一の符号を付して詳細な説明は省略する。即ち、図 5 の音声認識装置は、認識結果記憶部 21 が新たに設けられて

いる他は、図 2 の音声認識装置と共通する。

認識結果記憶部 2 1 には、制御部 1 1 が出力する音声認識結果が供給されるようになっており、認識結果記憶部 2 1 は、その音声認識結果を記憶する。なお、認識結果記憶部 2 1 は、例えば、1 つ又は複数の音声認識結果を記憶することができる分の記憶容量を有しており、また、その記憶容量の分だけ音声認識結果を記憶すると、次に供給される音声認識結果を、最も古い記憶値に上書きする形で記憶するようになっている。従って、認識結果記憶部 2 1 では、最新の音声認識結果の履歴が記憶される。そして、認識結果記憶部 2 1 に記憶された最新の音声認識結果の履歴は、単語予備選択部 1 3 が参照することができるようになっている。

次に、図 6 のフローチャートを参照して、図 5 の音声認識装置による音声認識処理について説明する。

図 5 の音声認識装置は、ステップ S 1 1 乃至 S 1 7 において、基本的に、図 4 のステップ S 1 乃至 S 7 における場合とそれぞれ同様の処理を行う。

ステップ S 1 7 において、制御部 1 1 が、最終スコアが最も大きいパスを構成するアークに対応する単語列をユーザの発話に対する音声認識結果として出力した後は、ステップ S 1 8 に進み、制御部 1 1 は、その音声認識結果を認識結果記憶部 2 1 に供給して記憶させて処理を終了する。

但し、ステップ S 1 4 では、単語予備選択部 1 3 は、次のような単語予備選択処理を行う。

即ち、単語予備選択部 1 3 は、図 4 のステップ S 4 における場合と同様に、注目ノードが有する時刻情報における開始時刻以降の必要な特徴量の系列を特徴量記憶部 1 2 から読み出し、辞書データベース 1 8 A の単語辞書に登録された各単語について、特徴量記憶部 1 2 から読み出した特徴量の系列を用いて音響スコアを計算するとともに、文法データベース 1 9 A に記憶された文法規則に基づいて言語スコアを計算し、その音響スコア及び言語スコアを総合評価した単語スコア（第 1 の尺度）を求める。単語予備選択部 1 3 は、単語スコアが上位 L 位以内の単語をマッチング処理の対象とする単語としてマッチング部 1 4 に供給する。

単語予備選択部 1 3 は、辞書データベース 1 8 A に登録されている単語のうち

の1以上を認識結果記憶部21に記憶されているという尺度（第2の尺度）に基づいて選択してマッチング部14に供給する。

ユーザが過去に行った発話に含まれる単語は、経験的に、その後の発話において再度発話されることが多い。そこで、図5の単語予備選択部13は、認識結果記憶部21に記憶されている音声認識結果を構成する単語と同一の単語を、辞書データベース18Aから特定単語として選択し、マッチング処理の対象とする単語としてマッチング部14に供給する。

マッチング部14は、単語予備選択部13から単語スコアに基づいて選択された1個の単語と、単語スコアに関係ない所定の条件（尺度）（ここでは、認識結果記憶部21に記憶されているという尺度）に基づいて選択された特定単語を受信すると、ステップS15において、それらの単語を対象としてマッチング処理を行う。

この場合も、単語予備選択部13において、音響スコアとは関係ない尺度に基づいて単語が選択されるので、例えば、環境雑音や回線雑音等に起因して音声の特徴量が大きな揺らぎを有し、そのような特徴量から計算される音響スコアを反映した単語スコアによっては選択されない単語であっても、精度の高い音響モデルを用いずに、マッチング処理の対象とされることになり、少ないリソースで音声認識精度を向上させることができる。

なお、音声認識装置には、図5に点線で示すように、ユーザが、音声認識結果を修正するための入力を与えるときに操作するキーボード等の入力部22をさらに設けることが可能である。この場合、制御部11が出力した音声認識結果に対してユーザが入力部22を操作することにより、その修正を行ったときには、制御部11には、自身が出力した音声認識結果に替えて（あるいは、その音声認識結果とともに）、その修正された音声認識結果を認識結果記憶部21に記憶させるようにすることができる。この場合においても、上述した場合と同様に、少ないリソースで音声認識精度を向上させることができる。

上述した一連の処理は、ハードウェアにより行うこともでき、あるいは、ソフトウェアにより行うこともできる。一連の処理をソフトウェアによって行う場合には、そのソフトウェアを構成するプログラムが汎用のコンピュータ等にインス

トールされる。

図7は、上述した一連の処理を実行するプログラムがインストールされるコンピュータの一例を示している。

プログラムは、コンピュータに内蔵されている記録媒体としてのハードディスク105やROM103に予め記録しておくことができる。

プログラムは、フロッピーディスク、CD-ROM(Compact Disc Read Only Memory)、MO(Magneto optical)ディスク、DVD(Digital Versatile Disc)、磁気ディスク、半導体メモリなどのリムーバブル記録媒体111に、一時的あるいは永続的に格納(記録)しておくことができる。このようなリムーバブル記録媒体111は、いわゆるパッケージソフトウェアとして提供することができる。

なお、プログラムは、上述したようなリムーバブル記録媒体111からコンピュータにインストールする他、ダウンロードサイトから、デジタル衛星放送用の人工衛星を介して、コンピュータに無線で転送したり、LAN(Local Area Network)、インターネットといったネットワークを介して、コンピュータに有線で転送し、コンピュータでは、そのようにして転送されてくるプログラムを通信部108で受信し、内蔵するハードディスク105にインストールすることができる。

コンピュータは、CPU(Central Processing Unit)102を内蔵している。CPU102には、バス101を介して、入出力インタフェース110が接続されており、CPU102は、入出力インタフェース110を介して、ユーザによって、キーボードや、マウス、マイクロホン等で構成される入力部107が操作等されることにより指令が入力されると、それにしたがって、ROM(Read Only Memory)103に格納されているプログラムを実行する。あるいは、また、CPU102は、ハードディスク105に格納されているプログラム、衛星若しくはネットワークから転送され、通信部108で受信されてハードディスク105にインストールされたプログラム、又はドライブ109に装着されたリムーバブル記録媒体111から読み出されてハードディスク105にインストールされたプログラムを、RAM(Random Access Memory)104にロードして実行する。これにより、CPU102は、上述したフローチャートにしたがった処理、あるいは上述したプロッ

ク図の構成により行われる処理を行う。CPU 102は、その処理結果を、必要に応じて、例えば、入出力インタフェース 110を介して、LCD(Liquid Crystal Display)やスピーカ等で構成される出力部 106から出力、あるいは、通信部 108から送信、さらには、ハードディスク 105に記録等させる。

ここで、本発明において、コンピュータに各種の処理を行わせるためのプログラムを記述する処理ステップは、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はなく、並列的あるいは個別に実行される処理(例えば、並列処理あるいはオブジェクトによる処理)も含むものである。

また、プログラムは、1のコンピュータにより処理されるものであっても良いし、複数のコンピュータによって分散処理されるものであってもよい。さらに、プログラムは、遠方のコンピュータに転送されて実行されるものであってもよい。

上述のように、単語予備選択部 13において、辞書データベース 18Aに登録された単語群から、特徴量を用いて計算される単語スコアに基づいて単語を選択するとともに、単語スコアとは異なる尺度に基づいて発話時間の短い単語や、言語的に、直前に認識された単語に接続しやすい単語、過去の音声認識結果に含まれる単語をも選択し、その選択した単語をマッチング処理の対象とするようにしたので、単語予備選択処理にあたって、精度の高い音響モデルを用いなくても、所定の単語の選択漏れを防止することができ、その結果、必要なリソースの増加を極力抑えながら音声認識精度の劣化を防止することができる。

さらに、マッチング処理の対象とする単語を、単語スコアとは異なる尺度、即ち、音声の特徴量を用いて計算される音響スコアに関係がなく、音声の特徴量の影響を受けない尺度に基づいて選択するようにしたので、例えば、雑音環境下において、その雑音の影響を受けることなく、マッチングの対象とする単語を選択することができ、音声認識装置の雑音耐性を向上させることができる。

マッチング部 14でスコア計算の対象となる単語は、単語予備選択部 13においてあらかじめ選択されているから、マッチング部 14による各単語のスコア計算は、前述したような、音響スコアの計算の一部を共通化する木構造のネットワークを構成せずに、各単語ごとに独立して行うことができる。この場合、マッチング部 14において各単語についてのスコア計算を時分割で行うようにすること

により、その計算のために確保する必要があるメモリ容量を小さく抑えることができる。

マッチング部 14 によるスコア計算は、各単語ごとに、時間的に独立して行うことができ、スコア計算に要するメモリ容量を使い回すことにより、必要とするメモリ容量を小さく抑えることができる。

この場合、単語のスコア計算を開始するときに、その単語が、どの単語であるのかを同定することができるから、そのスコア計算の初期の段階で、言語スコアを適用して枝刈りを行うことが可能となり、その結果、前述したような、単語を同定することができないことによって無駄な計算が行われることを防止することができる。

なお、図 2 や図 5 に示した音声認識装置は、例えば、音声によってデータベースの検索を行う場合や、各種の機器の操作を行う場合、各機器へのデータ入力を行う場合、音声対話システム等に適用可能である。より具体的には、例えば、音声による地名の問合せに対して、対応する地図情報を表示するデータベース検索装置や、音声による命令に対して、荷物の仕分けを行う産業用ロボット、キーボードの代わりに音声入力によりテキスト作成を行うディクテーションシステム、ユーザとの会話を行うロボットにおける対話システム等に適用可能である。

また、単語予備選択部 13 において、音韻数や品詞に基づいて選択される特定単語は、他の単語とは区別して、単語辞書に登録しておくことが可能である。

さらに、上述の例では、単語予備選択部 13 において、音響スコア及び言語スコアを総合評価した単語スコアに基づいて L 個の単語を選択するようにしたが、L 個の単語は、その他、例えば、音響スコアだけに基づいて選択することも可能である。

また、マッチング処理の対象とする単語を音響スコアが反映された単語スコアとは異なる尺度に基づいて選択する場合においては、その尺度として、複数の尺度を用いることが可能であり、これにより、例えば、発話時間の短い単語と、過去の音声認識結果に含まれる単語の両方を、必ず、マッチング処理の対象とする単語とすること等が可能である。

産業上の利用可能性

本発明は、音声認識の対象とする単語群から、1以上の第1の単語が、特徴量を用いて計算される第1の尺度に基づいて選択されるとともに、1以上の第2の単語が、第1の尺度とは異なる第2の尺度に基づいて選択され、その選択された第1及び第2の単語について、スコアが計算されるので、第2の単語が第1の尺度に基づいて選択されないことによる音声認識精度を劣化を防止することができる。

請求の範囲

1. 入力された音声に対してその音声認識結果の音響的な尤度を反映したスコアを計算し、そのスコアに基づいて前記音声を認識する音声認識装置において、

前記音声の特徴量を抽出する抽出手段と、

音声認識の対象とする単語群から1以上の第1の単語を前記特徴量を用いて計算される第1の尺度に基づいて選択するとともに、1以上の第2の単語を前記第1の尺度とは異なる第2の尺度に基づいて選択する選択手段と、

前記選択手段において選択された前記第1及び第2の単語について前記スコアを計算するスコア計算手段と、

前記スコアに基づいて前記音声の音声認識結果としての単語列を確定する確定手段とを備える音声認識装置。

2. 前記選択手段は、音韻数を前記第2の尺度として音韻数が所定の条件を満たす単語を前記第2の単語として選択する請求の範囲第1項記載の音声認識装置。

3. 前記選択手段は、品詞を前記第2の尺度として品詞が所定の条件を満たす単語を前記第2の単語として選択する請求の範囲第1項記載の音声認識装置。

4. 前記選択手段は、言語的な尤度を前記第2の尺度として言語的な尤度が高い単語を前記第2の単語として選択する請求の範囲第1項記載の音声認識装置。

5. 上記装置は、さらに音声認識結果を記憶する記憶手段を備え、

前記選択手段は、前記記憶手段に記憶されていることを前記第2の尺度として前記記憶手段に記憶されている音声認識結果に含まれる単語を前記第2の単語として選択する請求の範囲第1項記載の音声認識装置。

6. 上記装置は、さらに音声認識結果を修正するための入力を与える入力手段をさらに備え、

前記記憶手段は、前記入力手段からの入力にしたがって修正された音声認識結果を記憶する請求の範囲第5項記載の音声認識装置。

7. 前記選択手段は、前記音声の特徴量を用いて前記スコアを計算し、そのスコアに基づいて前記第1の単語を選択する請求の範囲第1項記載の音声認識装置。

8. 入力された音声に対してその音声認識結果の音響的な尤度を反映したスコア

を計算し、そのスコアに基づいて前記音声を認識する音声認識方法において、
前記音声の特徴量を抽出する抽出ステップと、

音声認識の対象とする単語群から 1 以上の第 1 の単語を前記特徴量を用いて計算される第 1 の尺度に基づいて選択するとともに、1 以上の第 2 の単語を前記第 1 の尺度とは異なる第 2 の尺度に基づいて選択する選択ステップと、

前記選択ステップにおいて選択された前記第 1 及び第 2 の単語について前記スコアを計算するスコア計算ステップと、

前記スコアに基づいて前記音声の音声認識結果としての単語列を確定する確定ステップとを備える音声認識方法。

9. 入力された音声に対して、その音声認識結果の音響的な尤度を反映したスコアを計算し、そのスコアに基づいて前記音声を認識する音声認識処理をコンピュータに行わせるプログラムが記録されている記録媒体において、

上記プログラムは、

前記音声の特徴量を抽出する抽出ステップと、

音声認識の対象とする単語群から 1 以上の第 1 の単語を前記特徴量を用いて計算される第 1 の尺度に基づいて選択するとともに、1 以上の第 2 の単語を前記第 1 の尺度とは異なる第 2 の尺度に基づいて選択する選択ステップと、

前記選択ステップにおいて選択された前記第 1 及び第 2 の単語について前記スコアを計算するスコア計算ステップと、

前記スコアに基づいて前記音声の音声認識結果としての単語列を確定する確定ステップとを備えることを特徴とする記録媒体。

1/7

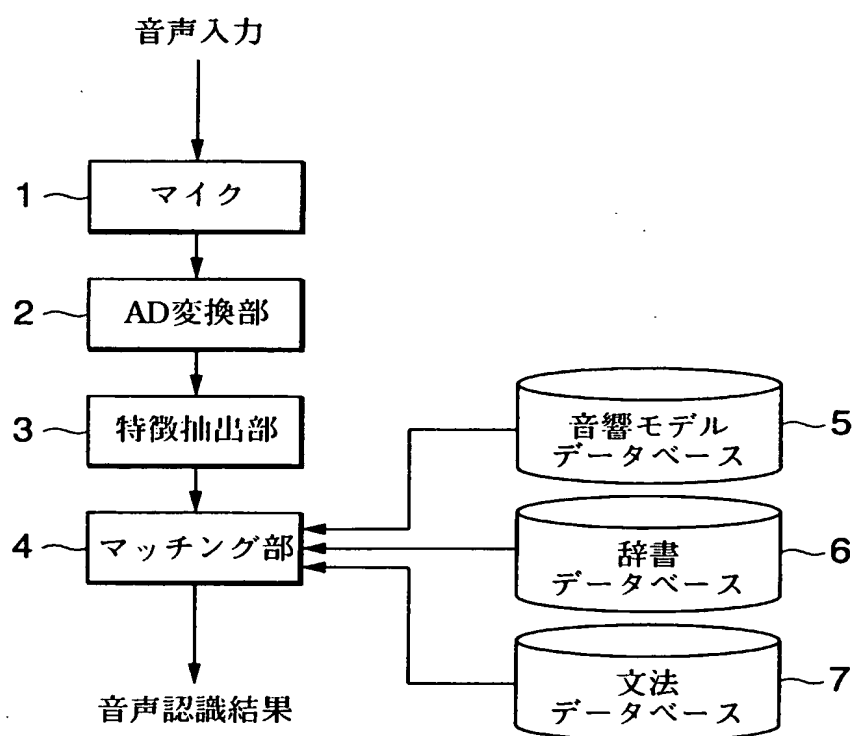


FIG.1

THIS PAGE BLANK (USPTO)

2/7

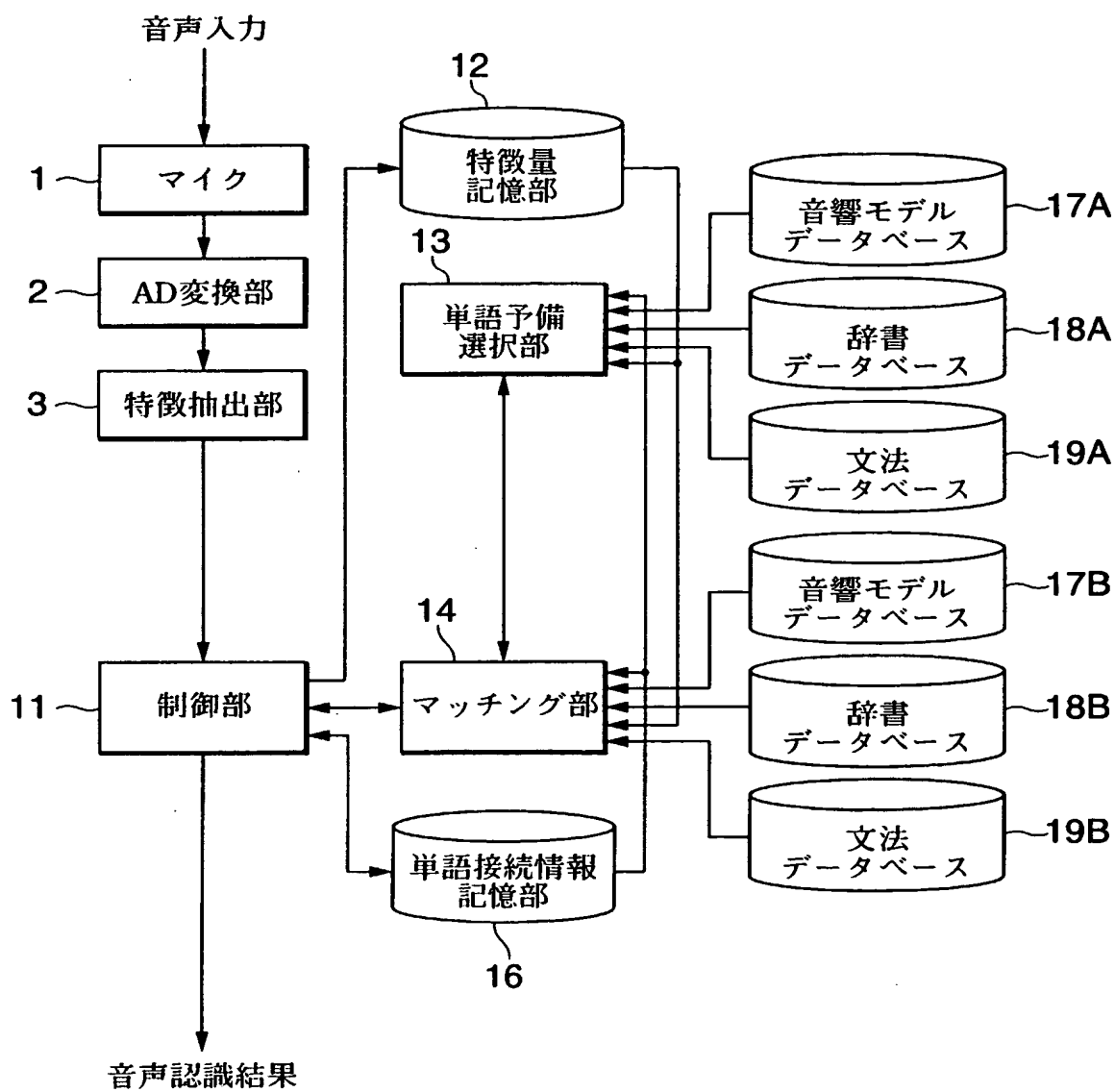


FIG.2

THIS PAGE BLANK (USPTO)

3/7

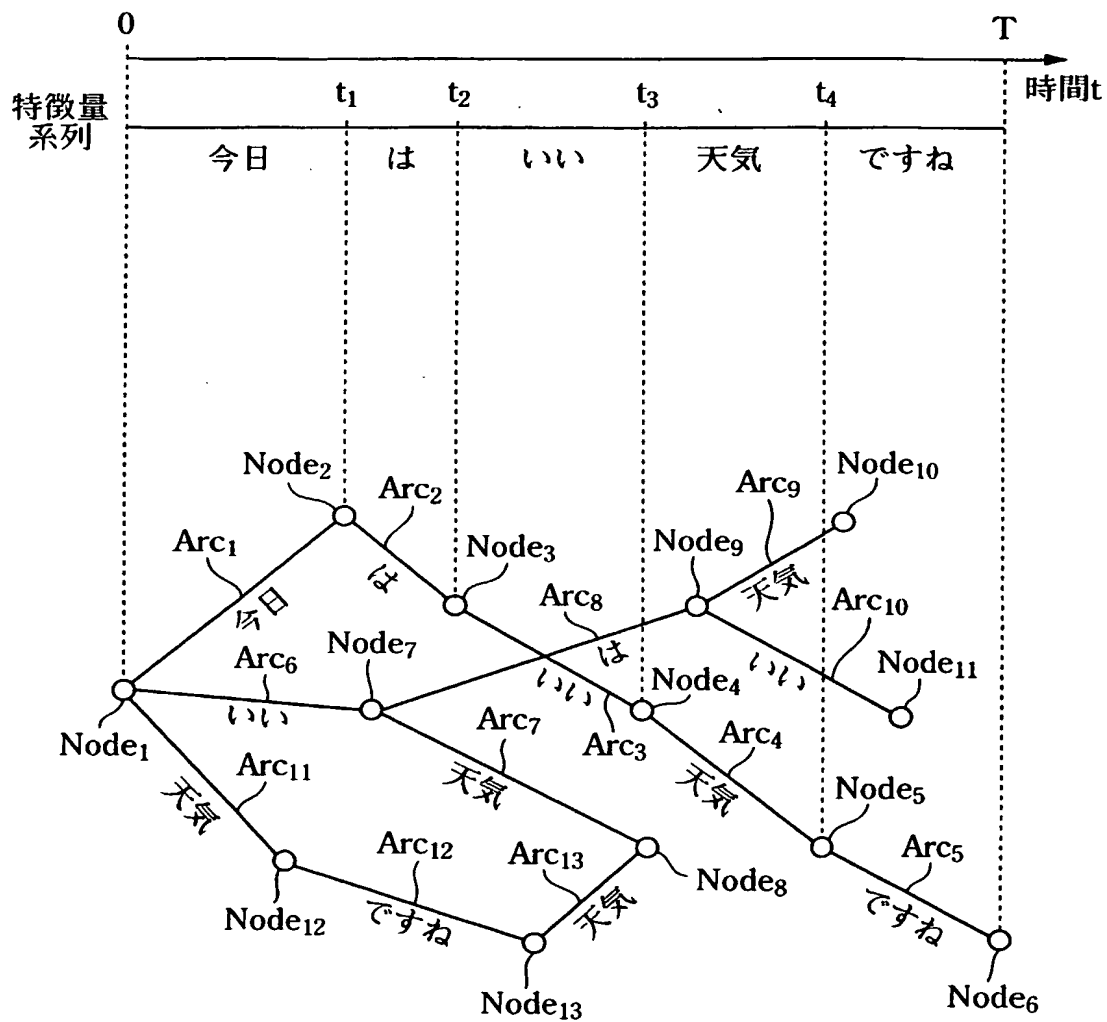


FIG.3

THIS PAGE BLANK (USPTO)

4/7

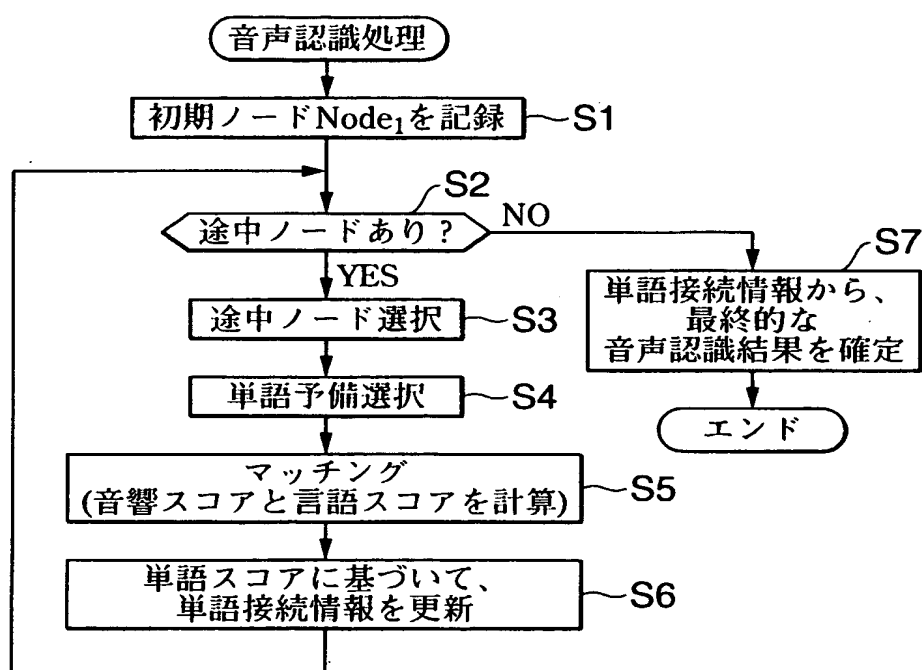


FIG.4

THIS PAGE BLANK (USPTO)

5/7

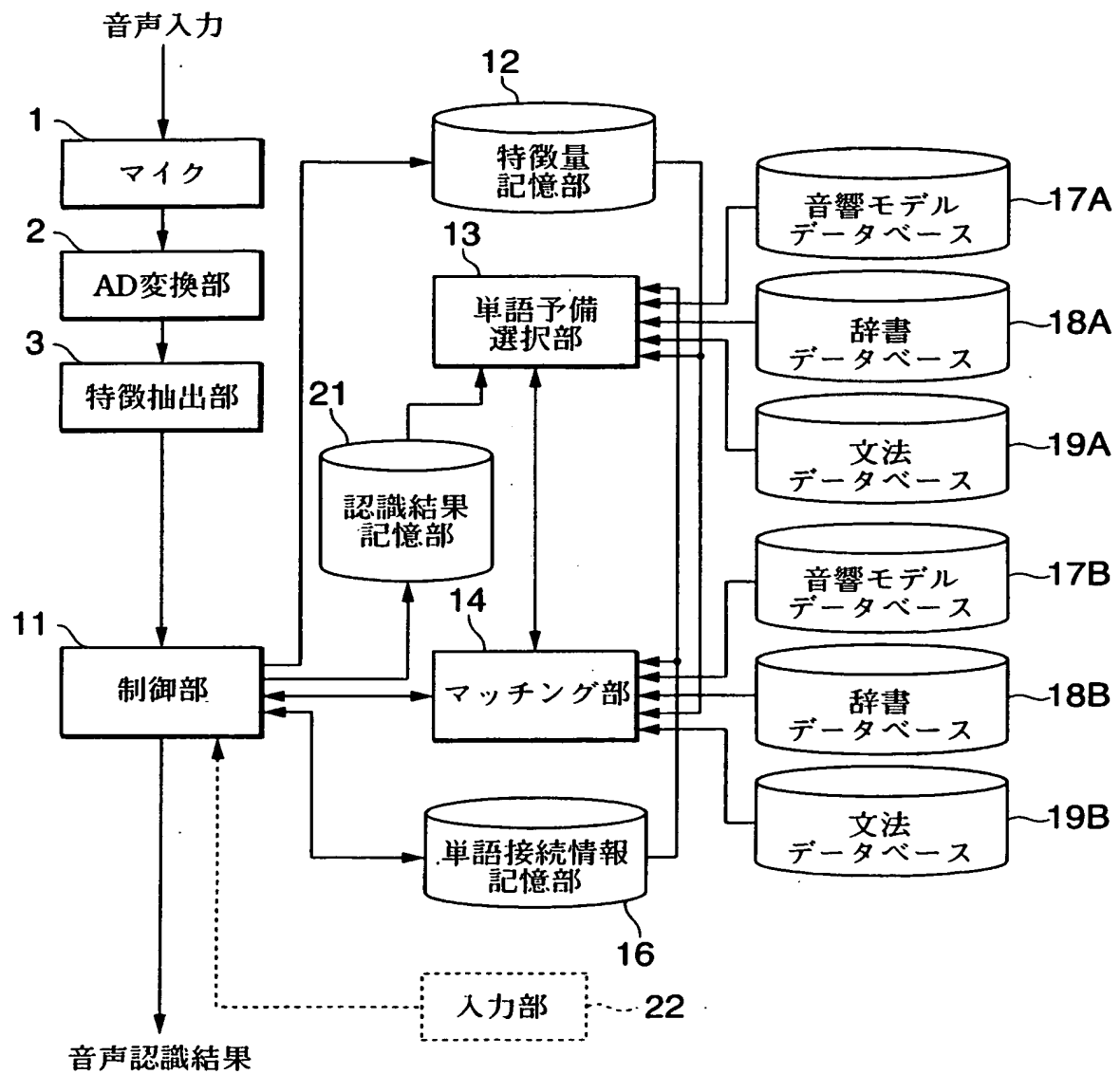


FIG.5

THIS PAGE BLANK (USPTO)

6/7

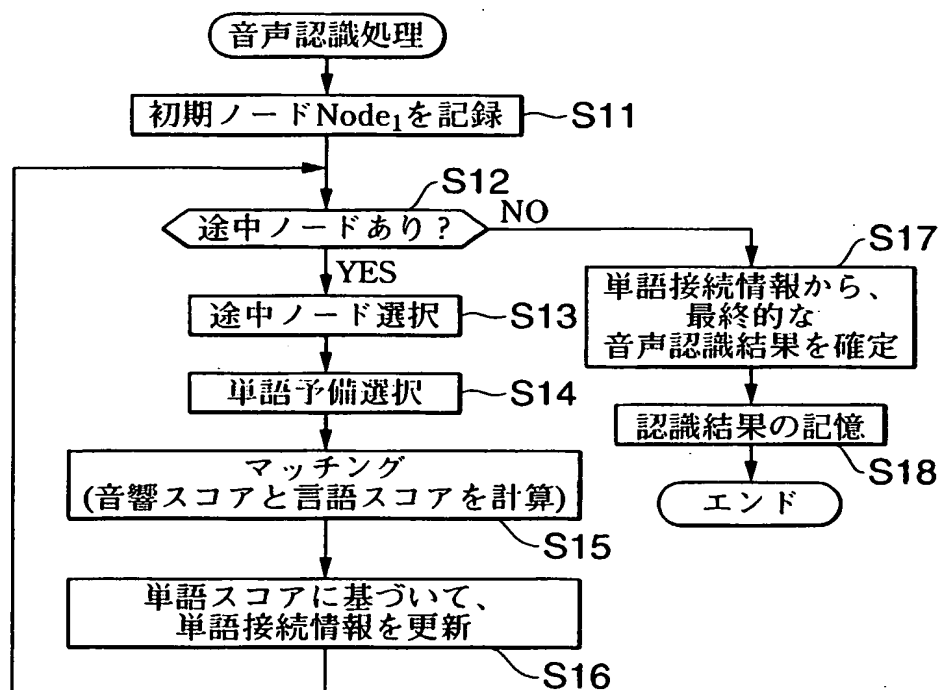


FIG.6

THIS PAGE BLANK (USPTO)

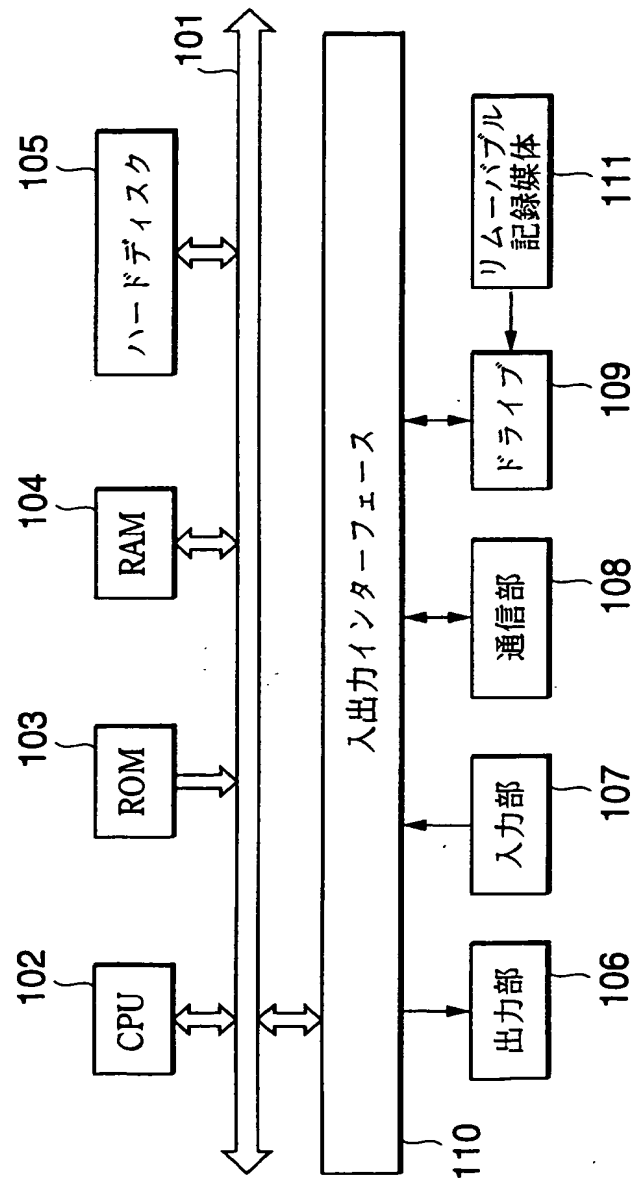


FIG.7

THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/01127

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl⁷ G10L15/28, 15/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G10L15/00-17/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2001
Kokai Jitsuyo Shinan Koho	1971-2001	Jitsuyo Shinan Toroku Koho	1996-2001

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JICST FILE (JOIS)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	IEEE Transactions on Speech and Audio Processing, Vol.1, No.1, January 1993, L. R. Bahl et al., "A Fast Approximate Acoustic Match for Large Vocabulary Speech Recognition", pp.59-67, ISSN1063-6676	1-9
A	JP, 57-86899, A (Mitsubishi Electric Corporation), 31 May, 1982 (31.05.82), Full text, all drawings (Family: none)	1-9
A	JP, 58-52696, A (Dainippon Printing Co., Ltd.), 28 March, 1983 (28.03.83), Full text, all drawings (Family: none)	1-9

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
29 March, 2001 (29.03.01)

Date of mailing of the international search report
10 April, 2001 (10.04.01)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

THIS PAGE BLANK (USPTO)

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int Cl⁷ G10L15/28, 15/18

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int Cl⁷ G10L15/00~17/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922~1996年

日本国公開実用新案公報 1971~2001年

日本国登録実用新案公報 1994~2001年

日本国実用新案登録公報 1996~2001年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICST科学技術文献ファイル (JOIS)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	IEEE Transactions on Speech and Audio Processing, Vol.1, No.1, January 1993, L.R.Bahl et al, "A Fast Approximate Acoustic Match for Large Vocabulary Speech Recognition", p.59-67, ISSN1063-6676	1~9
A	JP, 57-86899, A(三菱電機株式会社), 31.5月.1982(31.05.82), 全文全図, (ファミリーなし)	1~9
A	JP, 58-52696, A(大日本印刷株式会社), 28.3月.1983(28.03.83), 全文全図, (ファミリーなし)	1~9

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの

「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」 口頭による開示、使用、展示等に言及する文献

「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」 同一パテントファミリー文献

国際調査を完了した日

29.03.01

国際調査報告の発送日

10.04.01

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

松尾 淳

印

5C

8842

電話番号 03-3581-1101 内線 3540

THIS PAGE BLANK (USPTO)